The Dissertation Committee for Jinsuk Lee

certifies that this is the approved version of the following dissertation:

MOLECULAR STUDIES OF COTTON FIBER INITIATION

Committee:

Z. Jeffrey Chen, Supervisor

Stan Roux

Alan Lloyd

David M. Stelly

Tom Juenger

MOLECULAR STUDIES OF COTTON FIBER INITIATION

by

Jinsuk Lee, B.S.; M.S.

Dissertation

Presented to the Faculty of the Graduate School of

the University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2008

DEDICATION

To my parents, Seungkook Lee and Sungsun Kim, without whom none of this would have been possible. To my parents-in-law, Jaehwa Kim and Sookhee Kim, for their understanding and support. Although we are now several thousand miles away, their wishes and encourgament have been always with me.

To my husband, Sanghyun Kim, for his endless love and support.

ACKNOWLEDGEMENTS

I would like to give my appreciation to my advisor, Dr. Z. Jeffrey Chen. Without his support, advice, understanding and encouragement, I couldn't have completed any of this work.

I would like express my sincere gratitude to other committee members, Dr. David Stelly, Dr. Alan Lloyd, Dr. Stan Roux, and Dr. Tom Juenger, for their careful reading of my dissertation and their valuable guidance.

I wish to express my appreciation to Dr. Andrew Woodward, Dr. Lu Tian, and Dr. Zhongfu Ni who helped and encouraged me in both professional and personal manners whenever I had difficulties.

I would like to thank Erika Lackey for her great help and support in the lab. I am grateful to wonderful previous and fellow lab members, Dr. Hyeon-se Lee, Dr. Jianlin Wang, Dr. Wenxiang Gao, Dr. Sukhawn Yang, Dr. Osama Hassan, Meng Chen, Sheetal Rao, Ning Wei, Paulus Fong, Dr. Zhiguo Han, Dr. Jianxin Liu, Dr. Dr. Ruth Chang, David Pang, Eundeok Kim, and Misook Ha. I would like to thank you to Foo Cheung for his great help on gene annotation and data analysis.

Finishing up Ph.D. was one of the hardest things in my life. I'm proud of myself for being able to go through this tough time. Finally, my deepest gratitute goes to my husband and son, Sanghyun Kim and Juyeon Kim, for their love and support made it possible to complete my research work.

MOLECULAR STUDIES OF COTTON FIBER INITIATION

Publication No._____

Jinsuk Lee, Ph.D.

The University of Texas at Austin, 2008

Supervisor: Z. Jeffrey Chen

Cotton fiber development is a fundamental biological phenomenon. In spite of its economical importance, a large proportion of cotton fiber initiation is unknown. A naked seed mutant (*N1N1*) was compared with its isogenic lines of cotton (*Gossypium hirsutum*, TM-1) using a 70-mer oligonucleotide microarray that contained 1,536 features designed from a subset of cotton fiber ESTs. Statistical analysis and quantitative RT-PCR identified 23 "fiber-associated" genes. The annotation suggested that the temporal regulation of genes involved in transcriptional and translational regulation, signal transduction, and cell differentiation during early stages of fiber development. To get a large view of fiber initiation, a new cotton oligonucleotide microarray was developed

containing sequences from an ovule EST library from *Gossypium hirsutum* L. TM-1 immature ovules (GH_TMO), a set from Jonathan Wendel's lab at Iowa State University, and the pilot set of oligos used for previous study. Global gene expression studies were performed with microdissected fiber initials (or epidermis) and inner ovules to investigate fiber preferentially expressed genes. Laser capture microdissection and antisense RNA (aRNA) amplification allowed us to collect fiber initials (0 DPA and 2 DPA) or epidermal layers (-2 DPA) from whole ovule tissues. The gene expression profiles of fiber initials showed up-regulation of fiber proteins, myb transcription factors, and hormonal regulators as well as trichome related factors during fiber initiation. In each developmental stage, different sets of gene categories in molecular function or biological processes were over- or under-represented, suggesting temporal regulation of genes during fiber development. One of the possible "fiber associated genes" found in microarray analyses, *RD22 like* gene (*GhRDL*), was highly enriched in the epidermis of cotton ovules during fiber initiation. The function of *GhRDL* was studied with the *Arabidopsis* trichome system which shares many similarities with fiber development. Overexpression of 35S::*GhRDL* into *Arabidopsis thaliana* Columia-0 induced seed hairs (or seed trichomes) and *pRDL:GUS* was localized in *Arabidopsis* seeds. This suggests that *GhRDL* plays an important role in the seed trichome development and can be a key player in cell differentiation and fiber development.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Cotton is the one of most important crops in the world. It is the major source of natural fibers and cottonseed oil and can be used for innumerable commodities including textile production and paper currency. Cotton is grown in more than seventy countries. Cotton industry in the world represents a multimillion dollar enterprise today from raw fiber production to finished cotton textile product despite competition from man-made fibers (May and Lege, 1999).

Origin and History of Cotton

The history of cotton starts with the genus *Gossypium*, which arose about 10-20 million years ago (Wendel, 1989; Wendel and Albert, 1992). This genus includes more than 40 naturally occurring species in tropical and subtropical regions of world (Wendel, 1989; Smith and Cothren, 1999). A remarkable diversification in morphology and ecology accompanied the global radiation of *Gossypium* (Percival et al., 1999). The eight diploid genomic groups were recognized based on cytogenetic observations with interspecific hybrid viability and fertility (Percival et al., 1999). These diploids can be grouped into

---

This dissertation follows the format of The Plant Cell

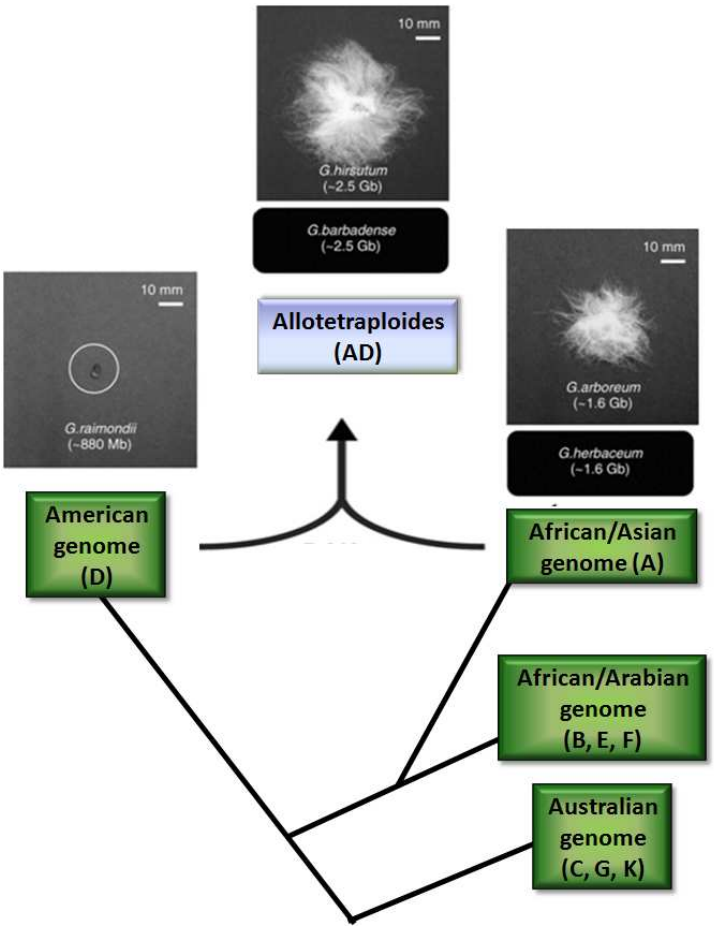

Figure 1.1. Evolution of allotetraploid cotton and cotton fibers*. Allotetraploidization occurred naturally 1–2 Mya between a fiber-producing A-genome extant species and a fiber-poor D-genome extant species, generating AD allotetraploid species (Wendel, 1989; Percival et al., 1999; Wendel and Cronn, 2003). There are five allotetraploid species, and two of them, *G. hirsutum* and *G. barbadense*, provide >95% of the modern commercial cotton crop. The genome sizes in parentheses are based on published work (Hendrix and Stewart, 2005). Mb, Mega base pairs (10^6); Gb, giga base pairs (10^9).

*This Figure is reformatted from "Gene expression changes and early events in cotton fibre development" by Lee *et al*. (2007). Annals of Botany. 100(7):1391-401.

three major lineages corresponding to their geographical places: Australia(C, G, K genomes), the Americas (D genome), and Africa/Arabia/Asia (A, B, E, and F genomes) (Smith and Cothren, 1999) (Figure 1.1). The K genome species is distinctive within the Australian lineage having upright flowers and lintless seeds with a fleshy elaisome that facilitates in dispersal by ants (Percival et al., 1999). The American D genome species have a variety of terpenoid aldehydes in the foliage with gossypol being the dominant terpenoid aldehyde (Wendel and Albert, 1992). The African/Arabian/Asian lineage includes the A genome progenitor of the allopolyploids. *Gossypium herbaceum* (A1) and *G. arboreum* (A2) are still cultivated on a small scale and provide a germplasm pool for several agronomically desirable traits (Endrizzi et al., 1985).

An important event in cotton genome evolution was the spontaneous formation of tetraploid species. These allopolyploids contain one genome that is similar to those found in the Old World A-genome diploids and one genome similar to those found in the New World D-genome diploids, and have been subsequently selected and domesticated as modern cultivated cotton (Endrizzi et al., 1985; Wendel, 1989). It is believed that this polyploidization event occurred about 1.5 million years ago and the AADD allotetraploid diverged into five species found throughout the tropics of the New World, making a new allotetraploid lineage (Wendel, 1989; Cronn et al., 1999; Adams et al., 2003; Desai et al., 2006) (Figure 1-1). The chloroplast genome of all allopolyploids are from an Old World A genome, suggesting that the seed parent in the initial polyploidization event was an African or Asian A-genome taxon (Wendel, 1989). For the D genome donor, *G.*

raimondii is considered to be the closest living model to the allopolyploids (Endrizzi et al., 1985).

The word "cotton" refers to the four lint-bearing species of *Gossypium* that were domesticated independently in four isolated regions in the New World and Old World. Two Old World cotton species (2n = 26), *G. arboreum* and *herbaceum*, arose from the African-Arabian gene pool, and the other two species (2n = 52), *G. barbadense* (pima or 'Egyptian' cotton) and *G. hirsutum* (upland or American cotton) evolved in the New World.

The two Old World species, *G. arboretum* and *G. herbaceum*, are shrubs or subshrubs with lobed leaves and yellow flowers (also white to red to purple in G. *arboretum*). *Gossypium arboreum* is found mostly in Asia with limited occurrence in southern Arabia and northern and eastern coastal Africa. *G. herbaceum* is distributed in Africa and Arabia, and also extends into India and Iraq. These two species have similar flavonoids, seed proteins, and chloroplast genomes (Cherry et al., 1972; Wendel, 1989; Wendel and Albert, 1992) but are distinguished by number of epicalyx bract teeth, epicalyx bract shape, and boll shape (Fryxell, 1979, 1992). The genetic studies of allelic variation also indicate that *G. arboretum* and *G. herbaceum* are also distinct species (Wendel, 1989). Wendel et al. (1989) used isozyme analysis to measure genetic divergence between *G. arboretum* and *G. herbaceum*. It has been hypothesized that the two species must have diverged from each other circa 1.4 million years B.P., long before human interference.

Two New World species, *G. barbadense* and *G. hirsutum,* are robust shrubs (Fryxell, 1979, 1992). The center of morphological diversity for *G. barabadense* is South America with range extensions into Mesoamerica and Caribbean. *Gossypium hirsutum* is indigenous to Mesoamarica, but its current range includes the Caribbean, the northern part of South America, and some pacific islands. *Gossypium hirsutum* can be easily distinguished by flower color and size. At the cytogenetic level, *Gossypium hirsutum* and *Gossypium barbadense* differ in ribosomal DNA sequences (Wendel et al., 1995) and chloroplast genomes (Wendel and Albert, 1992).

All four of *Gossypium arboreum, Gossypium herbaceum, Gossypium hirsutum,* and *Gossypium barbadense* went through a parallel series of developments under human selection as each species was transformed from an occasional source of lint to an industrial source of textile fiber (Fryxell, 1979) . Even though the Old World cottons are still cultivated in some areas of Africa and Asia, they have been replaced by the agronomically superior New World cottons that now dominate world cotton production. *G. barbadense* (AD_2) cultivars that are grown in Egypt, India, and the United States have long, strong and fine fibers, but lower yield than *G. hirsutum* (AD_1) cultivars that account for more than 90% of the world's annual cotton crop (Smith and Cothren, 1999).

Cotton Fiber Development

Seed fibers are characteristic for the genus *Gossypium*. They are single, elongated cells derived from the growth and differentiation of the outer ovule epidermal cells at or near

the day of anthesis (Benedict, 1984). The functions of cotton fibers are unknown, but may be related to seed dispersal, protection against herbivores, and extreme climate conditions (Ryser, 1999). The seeds of all *Gossypium* species have epidermal hairs. The wild species have only one type of hair whereas the cultivated cotton species have two types of differentiated seed hairs: fuzz and lint (Fryxell, 1963). The fuzz hairs are short, cylindrical, tightly adherent, and morphologically similar to the seed hairs of the wild species. In contrast, the lint hairs are longer and have thinner secondary cell walls with cellulose microfibrils, and also develop in different manner. Fuzz hairs do not begin elongation until 5-10 days after anthesis, but lint hairs begin to elongate on the day of anthesis (Mauney, 1984). Among the extant diploids resembling the presumed ancestors of tetraploid cotton, the AA genome species are fiber-producing, but the DD genome species produce only very few lint fibers, not spinnable fibers (Percival et al., 1999; Smith and Cothren, 1999; Applequist et al., 2001). Interestingly, allotetraploid species produce more abundant and higher quality fibers than the extant descendants of its ancestral species (Lee et al., 2007).

In general, growth in plants is a combination of cell division and cell elongation. However, cotton fibers elongate as single epidermal cells without any complications from cell division. This is a unique feature of fiber cells because the neighboring cells in the epidermal layer are subjected to continuous cell division as the seed develops (Basra and Saha, 1999). It has been shown that genes controlling fiber cell division exist and the age of the ovule is a critical factor in fiber cell division (Van't Hof and Saha, 1997). Another

significant factor in cotton fiber development is the cell cycle phase during the initial stage of development. When the S-phase inhibitor 5-aminouracil (5-AU) was applied, 98% of cotton fiber cells in ovules at 2 DPA were arrested at G1 phase that controls cytokinesis, whereas fiber cells without 5-AU passed through the S phase and underwent cell division (Van't Hof and Saha, 1997).

Cotton fibers undergo four overlapping developmental stages: fiber cell initiation, elongation, cellulose biosynthesis, and maturation (Basra and Malik, 1984; Tiwari and Wilkins, 1995; Wilkins and Jernstedt, 1999). The signals important to fiber cell differentiation must occur prior to the formation of fiber cell initials (Lee et al., 2006). Primordial fiber cells that are destined to become lint fibers emerge at, or prior to, anthesis, where they continue to develop up to 5 days post-anthesis (DPA), and then proceed to cell elongation that makes the fiber initials 1000 to 3000 times longer in diameter. Other epidermal cells, which form another type of fiber called “fuzz” fiber, are initiated 4 to 10 DPA but do not attain lengths greater than 10 mm. The fuzz fibers are usually about 1.5 to 3.3 mm long and remain attached to the seed coat after ginning, giving the seed a fuzzy appearance (Basra and Saha, 1999). During the fiber elongation stage, the fibers are enclosed in a thin primary cell wall that is plastic enough to accommodate the incorporation of new cell wall material (Basra and Saha, 1999). This is the result of a complex interplay between cell turgor and cell wall extensibility. At this stage, the level of cellulose in these very thin walls is relatively constant and microfibril deposition parallels the roughly transverse orientation of the cortical microtubule network

(Seagull, 1990; Seagull, 1992a, 1992b). The primary wall biosynthesis commences at 5–20 DPA depending on the cotton species, cultivar, and environment. The fiber thickening is accomplished through cellulose biosynthesis which occurs at 15 - 25 DPA (Meinert and Delmer, 1977; Lee et al., 2006). The fiber secondary cell wall consists of nearly pure cellulose. The overlapping period of fiber elongation and cellulose biosynthesis seems to be a major determinant of fiber length that length being considerably longer in *G. barabdense* (pima cotton) compared to *G. hirsutum* (upland cotton) (Benedict et al., 1973; Schubert et al., 1973). Furthermore, the timing, extent, and pattern of deposition of cellulose have an influence on fiber quality. Once fibers enter the phase of active cellulose synthesis, fiber elongation eventually stop and the fiber at this stage becomes a factory dedicated to cellulose synthesis. Finally, fiber cells mature from 50 to 60 DPA when cotton bolls open (Basra and Malik, 1984). The "programmed cell death" occurrs during this final phase of fiber development, wherein specific nucleases and proteases are induced, interceullular contents are degraded, and cells eventually die.

Cotton Fiber Initiation

Fiber cell initiation process is rapid and quasi-synchronous. Cell fate determination undoubtedly precedes the formation of morphologically visible fiber cell initials (Lee et al., 2007). Only 15–25% of the cells differentiate into the commercially important and spinnable ''long'' fibers (Basra and Malik, 1984; Wilkins and Jernstedt, 1999; Kim and Triplett, 2001; Lee et al., 2006). Cotton fiber initials first become visible at the day of

flowering through bulging and spherical expansion at the crest of the funiculus near the chalazal end of the ovule (Ryser, 1999).

During the fiber initiation stage, the nucleoli are active in the formation of pre-ribosomal particles necessary for ribosome formation and protein synthesis (Berlin, 1986). A large proportion of the ribosomes used for rapid fiber elongation is probably produced at this time (Ryser, 1999). Fiber initials are characterized by an increased volume of the nucleus and an increased number of Golgi complexes which may be involved in the biosynthesis of cell wall polysaccharides (Berlin, 1986; Kosmidou-Dimitropoulou, 1986). The activity of β-glycerophosphatase has been demonstrated specifically in fiber initials but not in other epidermal cells or in the epidermal cells of a fiberless mutant, suggesting increased carbohydrate metabolism for fiber development (Joshi et al., 1985). Another enzyme of carbohydrate metabolism, *sucrose synthase* gene (*SuSy*), was also reported to play a role in carbohydrate partitioning and turgor pressure maintenance during rapid cell growth and expansion (Hendrix, 1990). At 1 DPA, SuSy protein is immunolocalized to the basal area of initiating fiber cells in wild type cotton, coincident with the high accumulation level of *SuSy* mRNAs. This localization pattern indicates a close relationship between *SuSy* and fiber cell differentiation (Nolte et al., 1995). Comparative analysis of a fiberless (*fl*, fiber cell defective) mutant and wild-type cotton shows that *SuSy* mRNA is abundant in the fiber cell initials of normal ovules relative to that of mutant ovules. This can be correlated with the defective fiber cell initials in the fiberless mutant (Ruan and Chourey, 1998). Furthermore, down-regulation of *SuSy* mRNA levels

in the ovular epidermis by RNAi is associated with adverse phenotypes, including impairment of fiber initiation, fiber elongation and seed development (Ruan et al., 2001; Ruan et al., 2003). This data suggests that sucrose synthases are essential for carbon partitioning and turgor pressure maintenance during rapid cell growth and expansion.

Immunofluorescence work has shown a reorientation of the cortical microtubules in fiber initials but not in other epidermal cells. This reorientation of the microtubules was not observed in the epidermis of a fiberless cotton cultivar, suggesting microtubule involvement in fiber initiation (Jernstedt et al., 1993).

Several lines of evidence indicate that fiber cell differentiation is a complex biological process involving many other pathways such as signal transduction and transcriptional regulation (Kim and Triplett, 2001; Li et al., 2002a; Ji et al., 2003; Arpat et al., 2004). *FIDDLEHEAD* (*FDH*) encodes an enzyme involved in the synthesis of long-chain lipids found in the cuticle (Lolle et al., 1992; Pruitt et al., 2000), and mutations in *FDH* suppress epidermal cell interactions in *Arabidopsis* and display a deleterious effect on trichome development (Yephremov et al., 1999). In cotton, the FDH-like gene is highly expressed in developing fibers (Li et al., 2002a).

Various studies on cotton fiber development have implicated plant hormones as crucial regulators in fiber development. The physiological effects of plant hormones on cotton fiber development have been examined with *in vitro* cultures of cotton ovules. For several decades, auxins and gibberellins, alone or in combination, have been known to promote fiber cell initiation in unfertilized ovules (Beasley and Ting, 1974). However,

unfertilized ovules cultured with an inhibitor of gibberellin biosynthesis were unable to initiate fibers even in the presence of an auxin (Sharma et al., 1995), indicating that both auxins and gibberellins are needed to promote fiber initiation. *In vivo* assays revealed a spike of auxin level in flower buds preceding fiber initiation, declining after anthesis, and rebounding after 4 DPA (Guinn and Brummett, 1988). Auxins also regulate abscission with abscisic acids (abscission-promoting hormone) influencing boll retention. The ratio of abscisic acid to IAA correlates closely with abscission frequency through boll development (Lee et al., 2007).

Recent studies using microarray analyses have uncovered a set of genes that may regulate fiber cell initiation. Microarrays provide a high throughput tool for studying temporal expression patterns of many genes during fiber cell development. Using a filter array containing 1536 cDNAs of cotton, Li et al. (2002a) compared expression patterns of the genes in the ovules at 5 DPA between the wild-type and a different *fl* mutant and ten genes identified were highly enriched in the wild-type cotton. These include the genes encoding a RESPONSIVE TO DEHYDRATION22 (RD22)-like protein (*GhRDL*), a putative acetyltransferase (*GhACY*), a FIDDLEHEAD homologue (*GhFDH*), a serine carboxypeptidase-like protein (*GhSCP*), two tubulin components (*GhTUA6* and *GhTUB1*) and the fiber protein E6 (*GhE6*) . In 2006 (Wu et al.), the microarray analysis of laser-captured tissues was used to identify up-regulated genes in the cotton fiber initials relative to epidermal cells, many of which encode putative proteins of cell membrane and primary cell wall and DNA metabolism.

Yang et al. (2006a) compared approx. 211,000 cotton ESTs derived from elongating fibers and non-fiber tissues with approx. 32,800 ESTs derived from an ovular EST library using an equal mixture of RNA isolated from -3, 0 and 3 DPA ovules of *G. hirsutum* 'Texas Marker-1' (TM-1) (Yang et al., 2006a). The comparative data revealed many additional genes potentially involved in complex biological networks led to fiber cell development. The genes encoding putative transcription factors, including MYB and WRKY family members, are enriched in early stages of fiber and ovule development. The data agree with the known roles of MYB and WRKY transcription factors in *Arabidopsis* leaf trichome development. The data have also shown that AA-subgenome-specific ESTs are selectively enriched in young ovules (Yang et al., 2006a) relative to DD-subgenome transcripts. Enrichment of AA-subgenome mRNAs in fiber-bearing ovules is consistent with the production of long lint fibers in AA-genome species and the absence of such fibers in DD species.

Cotton Seed Trichome and *Arabidopsis* Leaf Trichome

Cotton fibers are classified as seed trichomes, which share many similarities with *Arabidopsis* leaf trichomes (Table 1.1). The models learned from cell fate determination and elongation in *Arabidopsis* leaf trichomes may provide a framework for understanding fiber cell initiation and elongation in cotton (Wang et al., 2004). In spite of striking similarities between two cell types, *Arabidopsis* leaf trichomes are branched (Hülskamp et al., 1994; Marks, 1997; Hülskamp and Schnittger, 1998; Hülskamp, 2004), whereas

cotton seed fibers are linear cells that never branched. Therefore, putative cotton genes identified in the stages of branch formation and growth directionality may have different functions in cotton fiber development. Moreover, endo-reduplication is common in *Arabidopsis* leaf trichomes but has inconsistent results in cotton fibers (Van't Hof, 1999; Taliercio et al., 2005). Many leaf trichome mutants are available in *Arabidopsis* (Hülskamp, 2004) and have been extensively used for the study of cell fate determination (Marks, 1997).

Trichome formation in *Arabidopsis* leaves are initiated by a 'trichome promoting complex' that consists of *GLABROUS1 (GL1)*, *TRANSPARENT TESTA GLABRA1 (TTG1)* and *GLABRA3* (*GL3)* (Szymanski et al., 2000). The epidermal cells in *Arabidopsis* leaves containing this complex induce *GLABRA2* (*GL2)* expression and develop trichomes, whereas the neighboring cells lacking this complex fail to initiate trichomes and become 'spacer' cells between trichomes (Szymanski et al., 2000; Hülskamp, 2004). The physiological changes in *Arabidopsis* leaf trichome development include DNA endoreduplication, rapid growth, and branch formation (Hülskamp, 2004; Schellmann and Hülskamp, 2005).

GL1 is a well-characterized MYB transcription factor required for leaf trichome initiation in *Arabidopsis* (Larkin et al., 1993). *GL1* is highly expressed in developing trichomes, and the gl1 mutation results in a trichome-less or glabrous phenotype (Larkin et al., 1993; Serna and Martin, 2006). *GL1* synergistically promote endoreduplication in

Table 1.1. *Arabidopsis* trichome related genes and their putative cotton orthologues

Developmental stages	*Arabidopsis* genes	Putative cotton orthologue	*E*-value*
Pattern formation	*GL1*	NP869199	1e-60
	Myb23	TC78581	7e-64
	TTG1	TC60200	1e-177
	GL3	TC71002, TC77343	1e-152, 1e-
	EGL3	TC71002, TC77343	1e-162, 1e-
	TRY	TC72522	6e-14
	CPC	TC72522	3e-15
	ETC1	TC80190	1e-11
	ETC2	TC72522	1e-11
	GL2	TC74707	0
Endoreduplication	*SIM*	N/A	N/A
	KAK	TC74666	9e-28
	SPY	CO122426	1e-129
	CPR5	TC71602	8e-13
	ICK/KRP	TC79148	5e-6
	RHL2	TC69621	e-130
	HYP	TC65874	4e-40
Branch formation	*AN*	TC71768	1e-164
	ST1	DN804424	4e-66
	FRC	TC76881	0
	TFCA	TC75687	7e-61
	TFCC	TC78431	1e-109
	ZWI	TC70950	0
Growth directionality	*GRL*	N/A	N/A
	KLK	TC62633	1e-15
	BRICK1	TC77933	7e-34
	ROP	TC67109	1e-114
	DIS1	DT561655	1e-105
	DIS2	TC68347	1e-134
	WRM	TC61308	0
	CRK	TC74982	0

N/A, Not available.

* *E*-value was estimated using tBLAST against the Cotton Gene Index CGI 8 (http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=cotton).

trichome cells (Szymanski and Marks, 1998). Overexpression of a cotton cDNA (L04497) encoding a putative MYB transcription factor in tobacco produces supernumerary epidermal trichomes on cotyledons and other organs (Payne et al., 1999), suggesting that cotton MYB transcription factors can influence leaf trichome initiation.

Expression studies of six MYB-related genes in *G. hirsutum* allotetraploids indicate that *GhMYB4* and *GhMYB6* display ovule-enriched expression patterns, and *GhMYB6* expression levels are high in fibers (Loguercio et al., 1999). Another gene, *GhMYB109* that encodes an R2R3 MYB-like transcription factor, is expressed specifically in fiber initials and elongating fibers (Suo et al., 2003). In fiber-bearing diploid species *G. arboreum* L. (AA), a transcript *GaMYB2* encoding a putative MYB transcription factor, is predominantly expressed early in fiber development (Wang et al., 2004). When *GaMYB2* was overexpressed in the *gl1* mutant, its glabrous phenotype was complemented. This suggests role for cotton MYB-like transcription factors in the development of leaf trichomes. The cotton MYB gene *GhMYB25* is expressed early during lint fiber initiation in cotton (Wu et al., 2006). *GhMYB25* is similar to Antirrhinum *AmMIXTA*, a putative R2R3 MYB protein. Overexpression of *GhMYB25* increases branches of leaf trichomes in transgenic tobacco, implying a link between cotton fiber initiation and leaf trichome development (Wu et al., 2006). Using laser-capture microdissection, Wu et al. (2006) further demonstrated that *MYB25* mRNA is enriched in the fiber initial cells relative to the non-fiber ovular epidermal cells.

In addition to *GL1*, trichome development in *Arabidopsis* leaves is mediated through several positive and negative regulators such as *TTG1*, *GL3*, *TRY*, *CAPRICE* (*CPC*), and *GL2* (Hülskamp, 2004; Schellmann and Hülskamp, 2005). *TTG1* and *GL3* are considered to be positive regulators along with *GL1* because mutation in any of these genes reduces the number of trichomes (Schellmann and Hülskamp, 2005). *GL3* has some functional redundancy with its close homologue *ENHANCER OF GLABRA3* (*EGL3*) (Zhang et al., 2003) and mutation in *GL3* alone causes smaller and less-branched trichomes (Payne et al., 2000). Sequence analyses have identified putative cotton orthologues to some of the *Arabidopsis* genes. Two cotton genes encoding *Arabidopsis* CPC orthologues were identified in the ESTs derived from fiber initials at 1 DPA and one of them was down-regulated in fiber initials at 1 DPA (Taliercio and Boykin, 2007) This suggests a negative role for putative cotton *CPC* orthologues in fiber development, which is reminiscent of the negative role of *Arabidopsis* CPC in the differentiation of leaf trichomes (Schellmann et al., 2002). The functions of these putative cotton genes in cotton fiber development should be tested further.

GL2 is expressed in trichomes and encodes a homeodomain transcription factor (Rerie et al., 1994; Pesch and Hülskamp, 2004). Genetic tests indicate that *GL2* acts downstream of *TTG1* and *GL1* because *gl2 gl1* and *gl2 ttg* double mutants lack trichomes, whereas plants with only the *gl2* mutation initiate trichomes normally (Hülskamp et al., 1994; Hülskamp, 2004). In *Arabidopsis*, *TTG1* functions in trichome formation and trichome spacing through lateral inhibition of differentiation in neighbouring epidermal

cells. Four tetraploid cotton genes homologous to *Arabidopsis TTG1* have been identified. The *Arabidopsis ttg1* mutant is rescued by a cotton orthologue, suggesting that equivalent *TTG1* functionality mediates both *Arabidopsis* leaf trichome and cotton fiber development (Humphries et al., 2005).

Outline of Dissertation

Despite its economical and biological importance, relatively little has been revealed about the nature and biological characteristic of cotton fiber initiation. Thus, the molecular study of fiber initiation in cotton is considered to be essential for uncovering genomic information to develop novel cultivars with favorable characteristics.

The ultimate goal of this research is to understand the developmental processes involved in the control of fiber development. Genetic and genomic approaches were used to investigate regulatory systems controlling the fundamental processes of fiber cell initiation and development. This study helps us to determine which genes control fiber initiation and development.

The study of morphological differences in fiber development among control and several mutant lines including *N1N1* and *n2n2* and among different developmental stages is described in Chapter II. The dominant mutation not only delayed the process of fiber cell formation and elongation but also reduced the total number of fiber cells, resulting in sparsely distributed short fibers.

Gene expression changes in TM-1 and *N1N1* mutant lines among four tissues were analyzed using spotted cotton oligo-gene microarrays which are also included in Chaper II. Using the *Arabidopsis* genes, we selected and designed approximately 1,334 70-mer oligos from a subset of cotton fiber ESTs. Quantitative RT-PCR analysis of 23 fiber-associated genes in seven tissues including ovules, fiber-bearing ovules, fibers, and non-fiber tissues in TM-1 and *N1N1* indicates a mode of temporal regulation of the genes involved in transcriptional and translational regulation, signal transduction, and cell differentiation during early stages of fiber development. Suppression of the fiber-associated genes in the mutant may suggest that the *N1N1* mutation disrupts temporal regulation of gene expression leading to a defective process of fiber cell elongation and development.

In Chapter III, differential gene expression was further studied with a large set of oligo-microarray which includes 22,787 oligosnucleotides on a single slide. To get only fiber initials or epidermis from whole ovule tissue, laser capture microdissection was applied and aRNA was amplified for hybridization. Ten different comparisons with epidermis and inner ovule tissues at 4 different developmental stages (-2 DPA, 0 DPA, 2 DPA, and 7 DPA) uncovered several fiber-associated genes which may play a crucial role during fiber initiation. Gene ontology classification suggests different set of genes are regulated in each developmental stage. Chapter III also describes how the candidate gene (*GhRDL*) is functioning in *Arabidopsis*. Interestingly, overexpression of *GhRDL* was able to induce seed trichomes, indicating that this gene may be involved in the developmental

process of seed trichome in *Arabidopsis*. This can provide evidence for an important role of *GhRDL* in cell differentiation and development during cotton fiber initiation.

CHAPTER II

GENE EXPRESSION STUDIES ON IMMATURE OVULES AND YOUNG FIBERS OF WILD TYPE COTTON AND NAKED SEED MUTANT (*N1N1*)

INTRODUCTION

Cotton fiber development is a fundamental biological phenomenon, yet the molecular basis of fiber cell initiation is poorly understood. We examined molecular and cellular events of fiber cell development in the naked seed mutant (*N1N1* and *n2n2*) and its isogenic line of cotton (*Gossypium hirsutum* L. cv. Texas Marker-1, TM-1).

Several ‘‘qualitative’’ mutants in fiber development have been reported (Kohel, 1973; Endrizzi et al., 1984). The best characterized of these are the naked seed loci, *N1N1* and *n2n2* which can produce “naked seed” phenotype. These mutants, discovered decades ago, lack most of the lint (long and spinable) fibers. The fuzz (short) fibers develop but eventually fall off the seeds to produce black or ‘‘naked’’ seeds. Phenotypically, *N1N1* is slightly more extreme and consistent than *n2n2*, which is more

---

This chapter is reformatted from “Developmental and gene expression analyses of a cotton naked seed mutant” by Lee *et el*. (2006). Planta. 223:418-432.

susceptible to changes in genotypes and environmental conditions. Interestingly, *N1N1* is dominant, whereas *n2n2* is recessive. In the recent studies of the fuzzless seed phenotypes in *Gossypium hirsutum*, the recessive form of the fuzzless seed allele *n3* was discovered (Turley and Kloth 2002). The dominant *N3* allele limits the expression of the recessive fuzzless seed phenotype and the segregating populations of homozygous *N3* locus had no significant effects on the expression of the *N1* fuzzless seed phenotype (Turley and Kloth 2002, 2008). Progeny analyses indicated that n3 is required for the expression of the recessive fuzzless seed allele n2 (Turley and Kloth 2002, 2008). Also, three monosomic lines with seed fuzz expression were all found in the D subgenome of allotetraploid cotton.

Gossypium hirsutum L. cv. TM-1 is an elite inbred line (>S40) derived from the obsolete cultivar ‘‘Deltapine 14’’. It has been widely used in research programs, isoline development, QTL analysis, and genetic and physical mapping. Mutant isogenic lines have been produced by backcrossing >6 generations to TM-1.

Observations using scanning electron microscope have shown that the dominant mutation not only delayed the process of fiber cell formation and elongation but also reduced the total number of fiber cells, resulting in sparsely distributed short fibers.

To gain a better understanding of gene regulation in the early stages of fiber development, gene expression changes in TM-1 and *N1N1* mutant lines among four tissues were analyzed using spotted cotton oligo-gene microarrays. Using BLAST analysis of *Arabidopsis* gene families, we selected and designed 1,334 70-mer oligos

from a subset of cotton fiber ESTs putatively encoding chromatin proteins, transcription factors, and proteins involved in cell wall biosynthesis and signal transduction pathways. We compared gene expression profiles in various tissues including young ovules (0 and 3 DPA) and non-fiber tissues in the naked seed mutant (*N1N1*) and its isogenic wild type (*Gossypium hirsutum* L. cv. TX Maker-1 or TM-1). Statistical analysis of the microarray data indicates that the number of significantly differentially expressed genes was 856 in the leaves compared to the ovules (3 days post-anthesis, DPA), 632 in the petals relative to the ovules (3 DPA), and 91 in the ovules at 0 DPA compared to 3 DPA, all in TM-1. Moreover, 117 and 30 genes were expressed significantly different in the ovules at three and 0 DPA, respectively, between TM-1 and *N1N1*. A subset of 23 fiber-associated genes was studied in seven fiber- and non-fiber tissues. Quantitative RT-PCR analysis of 23 fiber-associated genes in seven tissues including ovules, fiber-bearing ovules, fibers, and non-fiber tissues in TM-1 and *N1N1* suggest that the expression of many downstream genes encoding transcriptional and translational factors, signal transduction proteins, and cell differentiation factors is affected in the *N1N1* mutant, which may lead to a defective process in the early stages of fiber cell elongation and development. Suppression of the fiber-associated genes in the mutant may suggest that the *N1N1* mutation disrupts temporal regulation of gene expression, leading to a defective process of fiber cell elongation and development.

RESULTS

To examine fiber cell differentiation in the *N1N1* mutant at the cellular level, we used scanning electron microscopy (SEM) to observe the development of fiber cell initials in the ovular surface during early stages (0–3 DPA) of fiber development (Figure 2.1). On the day of anthesis (0 DPA), ◎25–30% of the cells in TM-1 began producing fiber cell initials (buds), whereas the protodermal cells in the *N1N1* mutant remained unchanged. Fiber cell formation proceeded rapidly in TM-1 and at 1 DPA the ovular surface near the chalazhal end was covered with evenly distributed fiber cell initials (Figure 2-1a, c). In *N1N1*, the fiber cell initials did not emerge until 1 DPA (or >12–16 h post-anthesis, Figure 2.1b, d) and developed at a relatively low density. We estimated the number of fiber cell initials was approximately ◎ 20–30% of that in TM-1, or ◎5–8% of the total protodermal cells. The elongation process of fiber cells was also slow and abnormal in the *N1N1* mutant (Figure 2.1f, h) compared to TM-1 (Figure 2.1e, g). As a result, only short fibers were formed in the *N1N1* mutant, whereas long fibers developed in TM-1. At 2 DPA, the fiber cells of TM-1 synchronously elongated with a relatively uniform length (Figure 2.2c), whereas the fiber cells in the mutant developed asynchronously and were distributed unevenly over the ovular surface (Figure 2.2d). From the longitude view, fiber cells initiated from the chalazals to micropylar end and covered ◎1/2–3/4 of the seed surface by the end of 2 DPA (Figure 2.2c–d). After 3 DPA, TM-1 ovules were covered densely with normal elongating fibers (Figure 2.1). Compared to the fiber cells in TM-1,

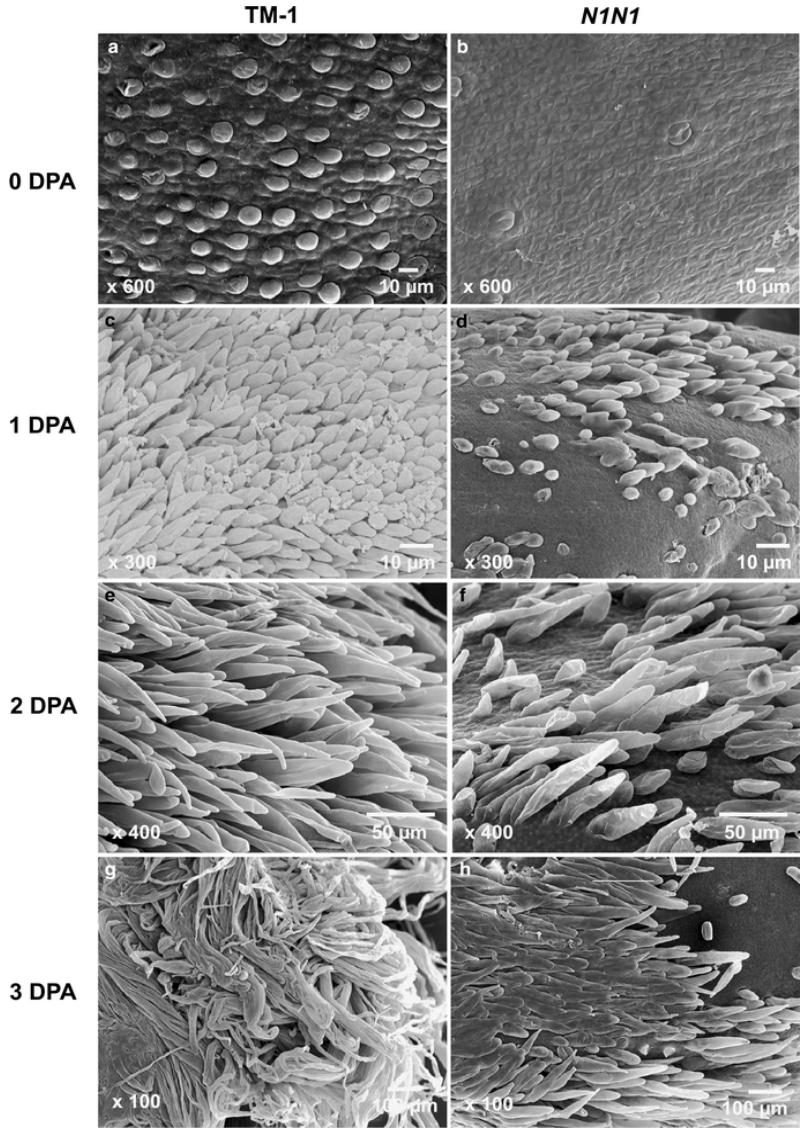

Figure 2.1. Delay of fiber cell initiation in the naked seed mutant (*N1N1*) observed using scanning electron microscopy (SEM). a–b Ovule surface of fiber cell initials at 0 DPA in TM-1 **(a)** and the isogenic mutant line, *N1N1***(b)**; c–d Fiber cells at 1 DPA in TM-1 **(c)** and *N1N1* mutant **(d)**; e–f Fiber cells in TM-1 **(e)** and *N1N1* **(f)** lines; and g–h Young fibers at 3 DPA in TM-1 **(g)** and *N1N1* mutant **(h)**.

the development of the sparsely dispersed fiber cells in the mutant was very slow, giving rise to short defective fibers. These short fibers eventually fell off the seed surface, and the seeds became "naked". The data indicate the *N1N1* mutation affects not only the number of protodermal cells for fiber differentiation but also fiber cell elongation.

Use of *Arabidopsis* genes to design cotton oligo-gene microarray

To understand the molecular basis of fiber cell differentiation and elongation, we developed spotted cotton oligo-gene microarrays using 16,695 EST Tentative Consensus (TC) assemblies (from the data posted in May 2002, http://www.tigr.org/tdb/tgi/plant.shtml), including ⊚10,000 from a diploid cultivated cotton, *G. arboreum* L. and ⊚7,000 from *G. hirsutum* L. In a comparison of nucleotide sequences, we found the cotton ESTs had the highest percentage of sequence identity with those of *Arabidopsis* and soybean. The sequence identity between cotton, rice, and maize ESTs was not significantly lower than that between cotton and other dicotyledonous plants (data not shown). This is because the majority of cotton ESTs was derived from the 5′ end of cDNAs. The 70-mer oligos selected showed a high percentage (60–85%) of sequence identity between cotton and other species. I selected 1,334 cotton ESTs encoding putative chromatin proteins, MYB, WRKY, MAP-kinase, cell wall and cell cycle proteins (Table 2.1) because these genes were predicted to play important roles in fiber cell differentiation and development

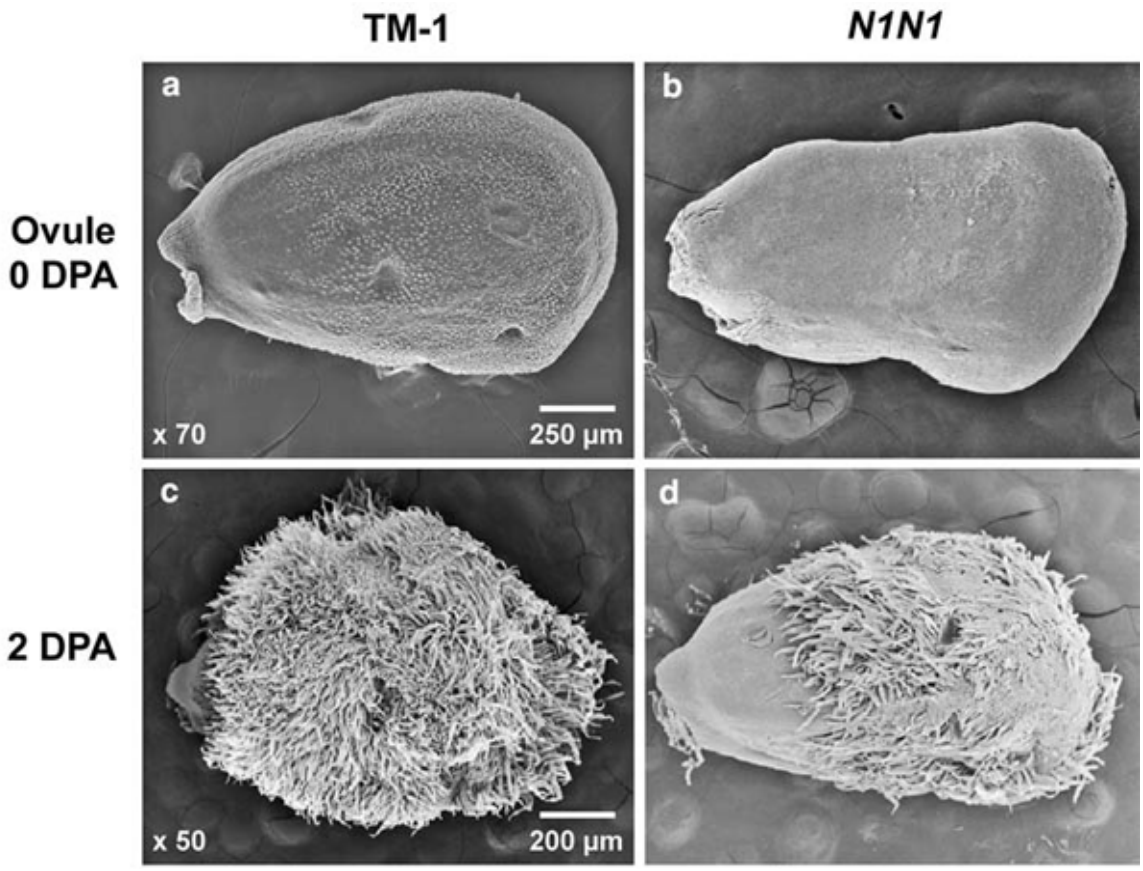

Figure 2.2. Ovule morphology of TM-1 **(a)** and its isogenic *N1N1* mutant **(b)** at 0 DPA; Distribution of fiber cells in TM-1 **(c)** and *N1N1* **(d)** at 2 DPA.

Table 2.1. *Arabidopsis* genes utilized for designing 70-mer oligos in cotton

Name of gene family	No. of *Arabidopsis* homolgs	No. of cotton ESTs matched
Chromatin factors	272	276
Myb transcription factors	131	56
WRKY transcription factors	(74) 71	99
Cell wall proteins	(56) 30	26
Cell cycle proteins	61	164
Cloned fiber-cDNAs	NA	130
Map kinases (MapK)	20	137
MapKK	10	143
MapKKK	80	228
Brassinosteriods	5	5
Drought and cold inducible genes	83	808
Subtotal	2,224-875 (overlap) = 1,334	
Controls	15	152
Chloroplast genes	(130) 121	NA
Mitochondrial genes	(67) 66	NA
Total	1,334 +15 + 121 + 66 = 1,536	

Numbers in parenthesis indicates the total number of gene families in *Arabidopsis*; NA: not applicable

(Arabidopsis Genome Initiative 2000) (Hülskamp, 2004). For the majority of protein groups, cotton has more homologous genes than *Arabidopsis*, suggesting that allotetraploid cotton contains many redundant genes resulting from polyploidization. An exception is that, excluding the overlaps, cotton fiber ESTs contain fewer numbers of the *Arabidopsis* MYB-related genes and chromatin genes, which is reminiscent of the estimate that cotton fiber genes represent 35% transcriptome in the cotton genome (Arpat et al., 2004). It is likely that some MYB and chromatin genes are under-represented in the current EST collections or expressed in the tissues such as immature ovules prior to fiber formation.

Microarray analysis of gene expression in fiber-bearing ovules and non-fiber tissues

We produced spotted cotton oligo-gene microarrays with a total of 1,536 elements using 1,334 cotton genes, 15 controls, and 121 chloroplast genes and 66 mitochondrial genes from *Arabidopsis* (Table 2.1). Four dye-swap hybridizations were performed as previously described (Tian et al., 2005), resulting in a total of eight replications for each experiment (Table 2.2). Five experimental comparisons (Table 2.2) were performed to analyze gene expression changes in the developing ovules containing young fiber cells (0 DPA, 3 DPA) in TM-1 relative to its isogenic *N1N1* mutant and in the ovules relative to the leaves or petals in TM-1 (Table 2.2). The microarray data were subjected to lowess normalization (Cleveland, 1979) and analyzed using a linear model.

Table 2.2. Microarray experimental design

Experiment	Slide	Dye-	RNA	Cy3	Cy5
1	1–4	1–2	RNA1	TM-1 (3 DPA)	TM-1 (Leaf)
	5–8	3–4	RNA2	TM-1 (Leaf)	TM-1 (3 DPA)
2	9–12	1–2	RNA1	TM-1 (3 DPA)	TM-1 (Petal)
	13–16	3–4	RNA2	TM-1 (Petal)	TM-1 (3 DPA)
3	17-20	1–2	RNA1	TM-1 (3 DPA)	TM-1 (0 DPA)
	21-24	3–4	RNA2	TM-1 (0 DPA)	TM-1 (3 DPA)
4	25–28	1–2	RNA1	TM-1 (3 DPA)	*N1N1* (3 DPA)
	29–32	3–4	RNA2	*N1N1* (3 DPA)	TM-1 (3 DPA)
5	33–36	1–2	RNA1	TM-1 (0 DPA)	*N1N1* (0 DPA)
	36–40	3–4	RNA2	*N1N1* (0 DPA)	TM-1 (0 DPA)

Table 2.3. The number of differentially expressed genes in fibers and non-fiber tissues and in TM-1 and *N1N1* that were detected using a common variance and/or a per-gene variance.

	Experiment	Number of genes				
		Per-gene variance	Common variance	Shared set*		
				Sum	Down-regulated gene[b]	Up-regulated genes[b]
1	TM-1 (3 DPA) vs. TM-1 (Leaf)	1037	909	856	412	444
2	TM-1 (3 DPA) vs. TM-1 (Petal)	880	673	632	326	306
3	TM-1 (3 DPA) vs. TM-1 (0 DPA)	449	162	91	8	83
4	TM-1 (3 DPA) vs. *N1N1* (3 DPA)	498	152	117	66	51
5	TM-1 (0 DPA) vs. *N1N1* (0 DPA)	195	65	30	20	10

[a] Shared: shared data sets of the significantly differentially expressed genes detected using both common variance and per-gene variance

[b] Up- and down-regulation referred to the genes in the former tissue in each comparison. For example, in the comparison of TM-1 (3 DPA) versus TM-1 (Leaf) the genes were down- or up-regulated in TM-1 (3 DPA)

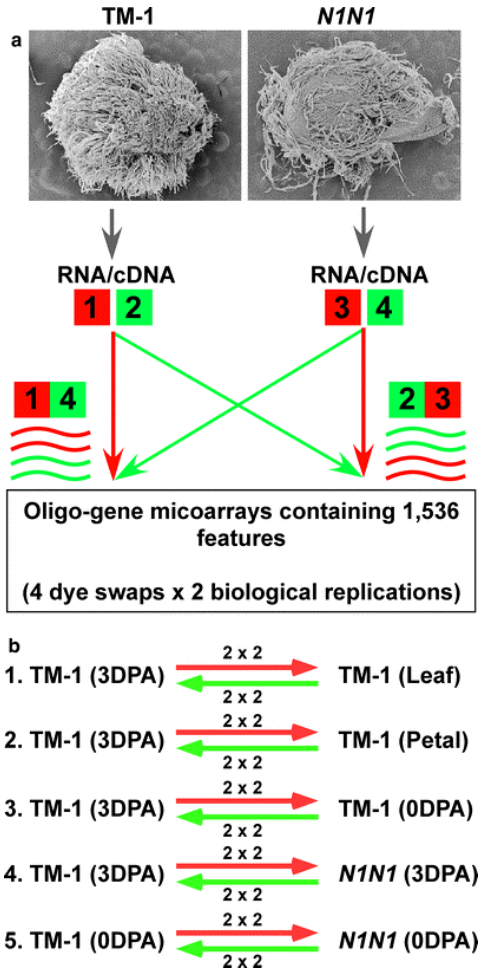

Figure 2.3. Experimental design of cotton oligo-gene microarray. **(a)** The RNAs are extracted from young fibers (3 DPA) of TM-1 and *N1* and reverse transcribed into cDNA. These cDNAs are labeled with florescent dye (either Cy3 or Cy5). Hybridization is performed using the probes prepared from Cy3- and Cy5- labeled cDNAs in TM-1 and *N1N1* (3 DPA). Each dye-swap consisted of 2 hybridizations, which was repeated four times making 8 replications of each experiment. **(b)** Repeated dye-swap experimental design for five sets of microarray comparisons. In each comparison, the dye-swap (2 hybridizations) as shown in A was repeated once (technical replication), and the repeated dye-swap (4 hybridizations) was performed using another RNA sample (biological replication), resulting in a total of 8-slide hybridizations with 4 technical replications (2 dye-swaps) and 2 biological replications (Table 2.2)

(Kerr and Churchill, 2001; Black and Doerge, 2002; Tian et al., 2005). The genes that displayed statistically significant expression differences in each experiment were selected using multiple comparison tests and a false discovery rate (FDR) at the 95% confidence level (Hochberg and Tamhane, 1987). The controls were not analyzed because no additional data normalization was needed in the linear model (Kerr and Churchill, 2001; Black and Doerge, 2002).

We compared gene expression changes in three sets of five experiments (Table 2.3, Figure 2.3b) using dye-swaps (Figure 2.3a). First, we compared gene expression divergence between fiber-bearing ovules and non-fiber tissues in TM-1. Although the ESTs were primarily derived from cotton fibers collected at 6–10 DPA, the majority of these genes were expressed in both fiber-bearing ovules and non-fiber tissues. A total of 856 (56%) genes were different expressed between the fiber-bearing ovules (3 DPA) and seedling leaves, whereas 632 (41%) genes were differentially expressed between the ovules and petals. Among the 856 genes that were differentially expressed in the leaves and ovules, 444 were up-regulated in the fiber-bearing ovules and 412 genes were up-regulated in the leaves. Excluding the chloroplast genes that are often highly expressed in the leaves, ◎350 genes were up-regulated in the leaves. Moreover, among the 632 genes that were differentially expressed in the ovules and petals, nearly half of them were up-regulated in the petals. The data indicate that 40–50% of the ESTs derived from cotton fibers are expressed in leaves and petals.

Second, we compared gene expression divergence between early stages of fiber development. Among the 91 genes that were significantly different in the ovules at 0 and 3 DPA, eight were down-regulated and 83 were up-regulated in the fiber-bearing ovules (3 DPA), suggesting that gene activation is a mode of temporal regulation of the biological pathways during early stages of fiber development.

Third, we compared gene expression divergence between TM-1 and *N1N1* in the early stages of fiber development. In the ovules at 3 DPA, 117 genes were expressed significantly differently between TM-1 and *N1N1*. Among these, 51 were up-regulated in TM-1 and 66 were up-regulated in *N1N1*. In the ovules at 0 DPA, 20 genes were up-regulated in *N1N1*, whereas ten genes were down-regulated in *N1N1*. Fewer genes detected in the ovules at 0 DPA than at 3 DPA may have indicated a bias of current cotton EST collections that were primarily derived from the late stages (6 - 10 DPA) of fiber development (Arpat et al., 2004).

We further analyzed differentially regulated genes among three comparative experiments using Venn-diagrams (Figure 2.4). If the genes are important to early events of fiber development, they would be up-regulated in the ovules at 3 DPA compared to those in the leaves in TM-1 and in the ovules in *N1N1*. In the comparison of up-regulated genes between the ovules and leaves or petals (Figure 2.4a), 590 genes were significantly up-regulated in both experiments. Among them, 284 and 146 genes were up-regulated in the ovules compared to the leaves and petals, respectively, while 160 genes overlapped in

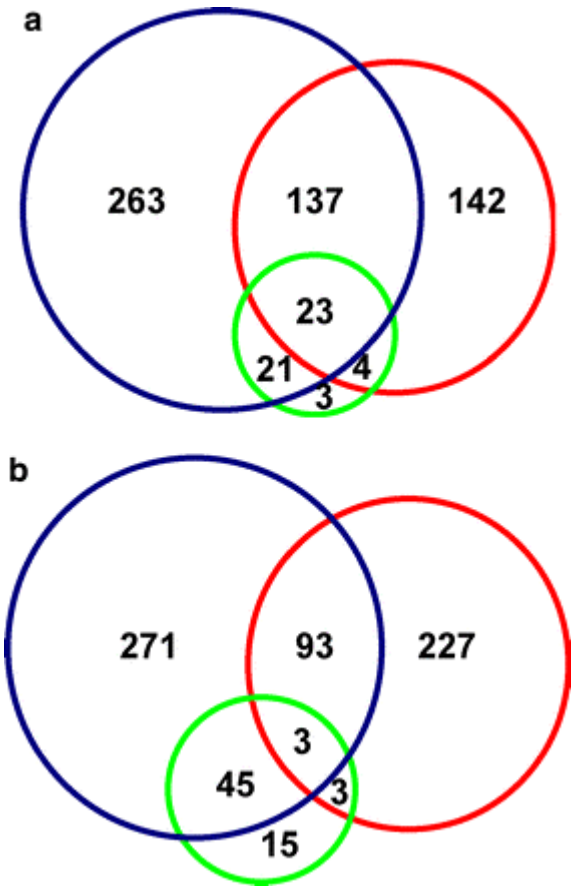

Figure 2.4. Comparative analysis of differentially expressed genes detected by microarray analysis. **(a)** Venn diagram displays the number of up-regulated genes in the ovules (3 DPA) compared to the leaves (blue) in TM-1, in the ovules (3 DPA) relative to the petals (red) in TM-1, and in the TM-1 ovules compared to *N1N1* ovules at 3 DPA (green); **(b)** Venn diagram displays the number of down-regulated genes compared between the ovules (3 DPA) and leaves (blue) in TM-1, between the ovules (3 DPA) and petals (red) in TM-1, and between the TM-1 and *N1N1* ovules at 3 DPA (green). The proportion shown in the diagrams may not be scaled.

the two comparisons, i.e., they were up-regulated in the ovules compared to both leaves and petals. Remarkably, among 51 genes that were up-regulated in the ovules at 3 DPA in TM-1 compared to *N1N1*, 48 (94%) matched the genes that displayed significantly higher levels of expression in the ovules relative to the leaves or petals. Among them, 23 (45%) were up-regulated in the ovules relative to both leaves and petals, whereas 21 (41%) were up-regulated in the ovules compared to the leaves, and four (8%) up-regulated in the ovules relative to the petals.

Interestingly, three genes (6%), two *HSP*-like genes (AI723022 and BG442970) and a stress-inducible gene (BF270275), were expressed differently only between TM-1 and *N1N1*, suggesting that *N1* mutation may be associated with stress responses that are not get determined.

Among 642 genes that were down-regulated in the ovules (3 DPA) compared to leaves and/or petals (Figure 2.4b), 316 (49%) were up-regulated in the leaves and 230 (36%) up-regulated in the petals, while 96 genes (15%) overlapped between the two comparisons. Among 66 genes that were down-regulated in the ovules (3 DPA) of TM-1 compared to *N1N1*, only three (5%) matched to the genes that were down-regulated in the TM-1 ovules compared to both leaves and petals, while 45 and three genes were down-regulated in the TM-1 ovules compared to the leaves and petals, respectively. The large number of the down-regulated genes detected in the ovules suggests that many ESTs derived from cotton fibers (3 DPA) were not highly expressed during the development of

Table 2.4. Primer sequences used for quantitative RT-PCR analysis of 23 genes in seven tissues

Target gene	Gene ID	Forward and reverse primer sequences (5' →3')
E6	BM356398	F:AAGAATGCCTACAAGTCCACTAAGC, R:CCTTGGTGCTCCATCCTTTC
FDH	AY072823	F: CGAGGCGTCGACTGTGATG, R:CTTGGGTCGGATCCTGGTT
RDL	AY072821	F: GCTTGCCTAGCGCTTGCT, R: TGGCAGCTTATAGCTCCAATATTG
WBC	TC27812	F: ACGATGGTTGCACCCTCTCT, R:CATTAGCTCGATTTGCGACATAGT
SAHH	TC27322	F: GAAACCAAAACTGGCATCATTG, R: GGGTGTCCAGTGGCACATC
STP	TC32989	F: CAGCCGCTTTGGTGAGTCA, R: CCAACCGGATGTTGCTTCA
NAP	TC39106	F: CGTCAATGAGCAACGAAGGA, R: GTGCTTCGGCGTTAAGAGCAT
SPP	TC27288	F: TGGATGCTTCAGCTCGGATA, R: CGACAACACATTCACTCCATCA
MYB25	TC38383	F: TCGATCCTGTCACCCACAAG, R:TTAGCAGCATCTTTAGGGTTACCA
RDL	AI726131	F: CGCAAATGAGTCGGAATGAAC, R: GGGTCTCCGACATTCACATGA
SAC25	TC37978	F: GGCGACGACCTGCTACGT, R:GTCTACAAAGTATTTCGCGGTGACT
CNX	AW561923	F: GCATTTTGCGCTTCTTGTCTT, R: TCAGCGGCAATGCAGAGA
EF1A	AW561930	F: TGGTGGTTTCGAAGCTGGTAT, R: GGGTAAACGCAAGCAAAGCA
HSP70	AW587470	F: TCGCCATGAACCCCTTGA, R: GCGGCATCGCTGAATCTAC
CEP52	BG442892	F: TCCCACCGGACCAACAAC, R: AGCCAAGGTACGGCCATCT
FAD3	TC32637	F: AAGCTACAAAAGCAGCGAAACC, R: AACGGAATCGGCCCTGAT
SRF6	TC28115	F: TCTCTCACGTTTTGCCGATGT, R: TGGGTGGCCGAAACTCA
PDF1	TC27374	F: GATGGAGAGGCAAAGAAGCAA, R: TGGGAGAACAAGGCAGCAA
BDC1	TC37094	F: CCCGAATGTGGCTCTTTTCTT, R: TTTCAATGAAATGCAGGCTCAT
BDC2	TC27702	F: GCTAGCCGAGGTGAGACGAA, R: CGATGTGGCGAAGTCAATCA
NOD26	TC38154	F: GACACCGGCGGGTTTAGTAG, R: CGAAACCGCCACAAACAAC
ACT	TC37642	F: CCACTTGTCTGCGACAATGG, R: GAACACAGCCCTTGGAGCAT
EF1-like	TC27277	F: CAGGGCAGATTGGAAATGGA, R: GCAAACTTGACCGCAATGTG

Table 2.5. List of 23 genes that were up-regulated in the fiber-bearing ovules (3DPA) in TM-1

GeneBank Accession	TIGR CGI[a]	Symbol	Description	TM-1 (3DPA)/ N1N1 (3DPA)[b]	TM-1 (3DPA)/ TM-1 (Petal)[b]	TM-1 (3DPA) /TM-1 (leaf)$_b$	Expression patterns in qRT-PCR analysis[c]
BM356398	TC27365	E6	Fiber protein E6 (clone CKE6-4A)	1.74 (2.07±0.23)	51.24 (28.20±3.51)	47.05 (168.66±11 .42)	III
AY072823	TC37712	FDH	Fiddlehead-like protein	1.94 (1.84±0.40)	2.38 (2.71±0.51)	2.69 (9.26±1.45)	IV
AY072821	TC37267	RDL1	Dehydration -induced protein RD22-like protein	1.82 (9.28±1.98)	1.63 (3.49±0.16)	5.35(229.32 ±22.65)	III
AW587505	TC27812	WBC1	ABC transporter	2.16 (5.37±0.20)	1.71 (1.10±0.10)	2.96 (8.83 ±1.42)	III
AW587516	TC27322	*SAHH*	S-adenosylho mocysteine hydrolase	2.12 (9.91±0.63)	44.07 (48.09±8.47)	12.23 (12.91 ±1.20)	III
AY094465	TC32989	*STP*	Putative sugar transport protein (*Arabidopsis thaliana*)	1.82 (3.68±0.24)	2.33 (2.66±0.10)	3.59 (21.14±1.4 0)	III
U12858	TC39106	*NAP*	Nucleosome assembly protein I-like protein	1.70 (2.05±0.04)	3.91 (4.55±0.85)	3.23 (12.15±0.1 7)	IV
BM356400	TC27288	*SPP*	Serine protease-like protein (*Arabidopsis thaliana*)	1.72 (3.98±0.43)	4.90 (107.88±8.94)	3.81(10.51 ±1.20)	III
AF336283	TC38383	*MYB25*	GhMYB25	1.86 (3.40±0.37)	3.59 (465.23±28.25)	3.02 (2307.50±1 89.94)	II
AI726131		*RDL2*	Dehydration -responsive protein RD22	2.32 (7.52±0.71)	2.76 (2.43±0.24)	13.94 (244.08±41 .08)	V
AY086643	TC37978	*SAC25*	Putative pod-specific dehydrogena se SAC25	2.71 (12.13±0.47)	3.16 (1637.62±63.5 1)	2.63(5.07± 0.22)	II
AW561923		*CNX*	Calnexin (*Glycine max*)	1.69 (4.31±0.63)	1.87(31.37 ±1.55)	5.26 (144.83±14 .73)	III
AW561930	TC27269	*EF1A*	Elongation	1.65	4.66(2.08±0.4	4.61(5.18±	IV

Table 2.5, cont.

			factor-1 alpha	(2.64±0.48)	3)	1.54)	
AW587470	TC39685	*HSP70*	Heat shock protein 70 (*Petunia x hybrida*)	1.80(3.45±0.58)	2.29(15.44±1.61)	4.62 (7.63±0.74)	IV
BG442892		*CEP52*	Ubiquitin/ribosomal protein CEP52 (*Trypanosoma brucei*)	1.58(5.81±0.30)	1.88(0.37±0.04)	2.18(1.90±0.27)	IV
BG447373	TC32637	*FAD3*	Omega-3 fatty acid desaturase, endoplasmic reticulum	1.70(5.05±0.44)	2.54 (93.50±6.60)	1.88 (45.36±4.47)	III
AY518291	TC28115	*SRF6*	Strubbelig recepter family 6 (*Arabidopsis thaliana*)	1.62(1.72±0.17)	3.09 (0.52±0.04)	3.21 (1.70±0.06)	IV
AF141375	TC27374	*PDF1*	Protodermal factor 1 (*Arabidopsis thaliana*)	1.70(2.64±0.21)	67.03 (636.96±31.32)	11.64 (24.68±2.11)	I
AY343972	TC37094	*BDC1*	BURP domain-containing protein	2.44(5.24±0.86)	3.56(2.23±0.22)	15.04(189.80±19.56)	III
AC004473	TC27702	*BDC2*	BURP domain-containing protein	1.79(19.67±2.53)	3.28(1161.27±154.42)	3.89(107.21±4.83)	III
L12258	TC38154	*NOD26*	Nodulin-26(*Glycine max*)	1.97(17.67±0.43)	2.31 (94.21±4.59)	5.89 (66.18±4.04)	III
AF059484	TC37642	*ACT*	Actin	1.70(5.26±0.16)	8.43(3.16±0.39)	7.05 (83.75±4.46)	V
AC026875	TC27277	*EF1-like*	Translation elongation factor	1.58 (3.87±0.53)	5.65 (7.88±1.19)	4.87 (10.66±1.01)	III

[a] Cotton gene index at TIGR (http://www.tigr.org/tigr-scripts/tgi/T_index.cgi?species= cotton)

[b] The upper and lower numbers in each column indicate the fold changes of gene expression in each comparison detected by microarray and qRT-PCR analyses, respectively

[c] See the text for the description of five gene expression patterns detected in qRT-PCR analysis

ovules and young fiber cells. The small number of overlapping genes that were down-regulated among the three comparisons indicates that up-regulation of gene expression in the *N1N1* ovules (3 DPA) may not be associated with the regular process of fiber cell development.

From these comparisons (Figure 2.5b), twenty-three genes (Table 2.5) that were up-regulated in the ovules at 3 DPA in TM-1 are the likely candidate genes important to early stages of fiber development. Indeed, many of them such as *E6* and *RDL1* are known to play roles in fiber development (John and Crow, 1992; John, 1996; Li et al., 2002a; Wang et al., 2004). The up-regulation also accounted for twenty other genes encoding nucleosome assembly protein, adenosylhomocysteinase, elongation factor (EF-1 alpha), calnexin homolog precursor, BURP domain-containing protein, ABC transporter, and sugar transporter, which are probably involved in the primary and secondary cell wall biosynthesis associated with rapid cell expansion and cellular elongation.

Up-regulation of the candidate genes is associated with fiber development

To study developmental regulation of fiber-associated genes detected by microarray analysis, we analyzed expression patterns of 23 genes (Table 2.5) in seven different tissues including leaves, petals, ovules at −3, 0, 3, and 5 DPA, and fibers at 10 DPA using quantitative RT-PCR (qRT-PCR) analysis. Data for six genes representing various expression patterns are shown in Figure 2.5a, and the rest of data in Table 2.5 and

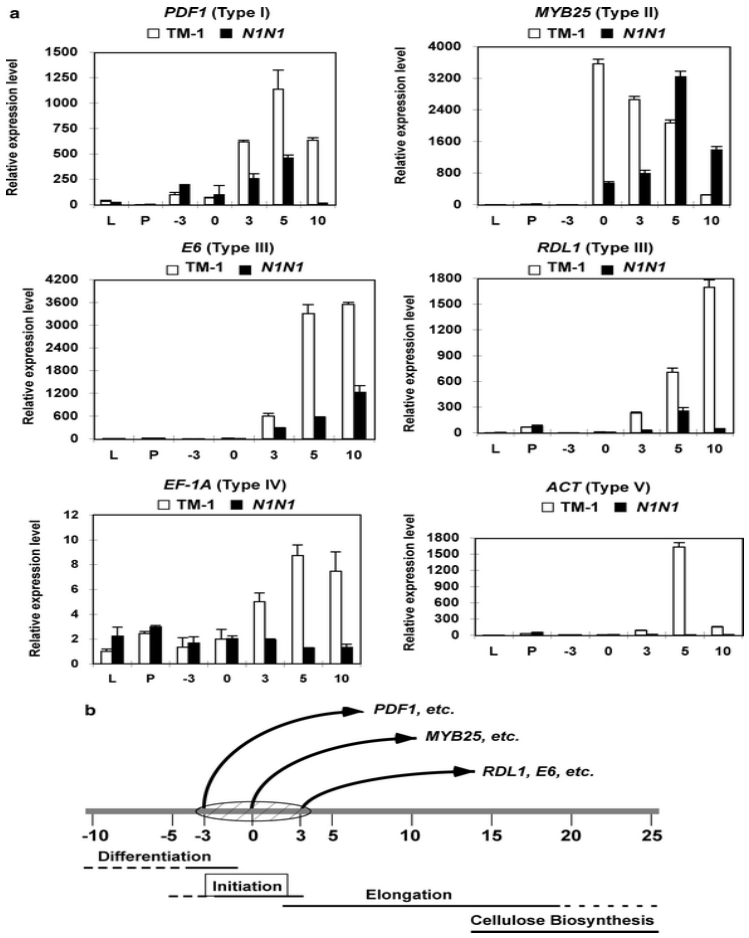

Figure 2.5. A sequential activation of the fiber-associated genes during early stages of fiber development. **(a)** Gene expression was analyzed in seven tissues including fibers and non-fibers using quantitative RT-PCR (qRT-PCR) analysis. *L*: leaves; *P*: petals; 0, 3, and 5: ovules at 0, 3, and 5 DPA, respectively; 10: fibers at 10 DPA. The expression patterns of six genes shown in the figure represented five gene-activation patterns (Type I–V) observed from early stages of fiber cell development. The expression patterns of other genes can be found in Table 2.5 and Figure 2.6 **(b)** A simplified model for temporal activation of gene expression during early stages of fiber cell development. The genes were up-regulated sequentially in the ovules at −3, 0, and 3 DPA during stages of fiber cell expansion and elongation. Down-regulation of the fiber-associated genes in the *N1N1* mutant led to a defective process of fiber cell development. The stage from −3 to 3 DPA (hatched oval) marked fiber cell initiation or formation.

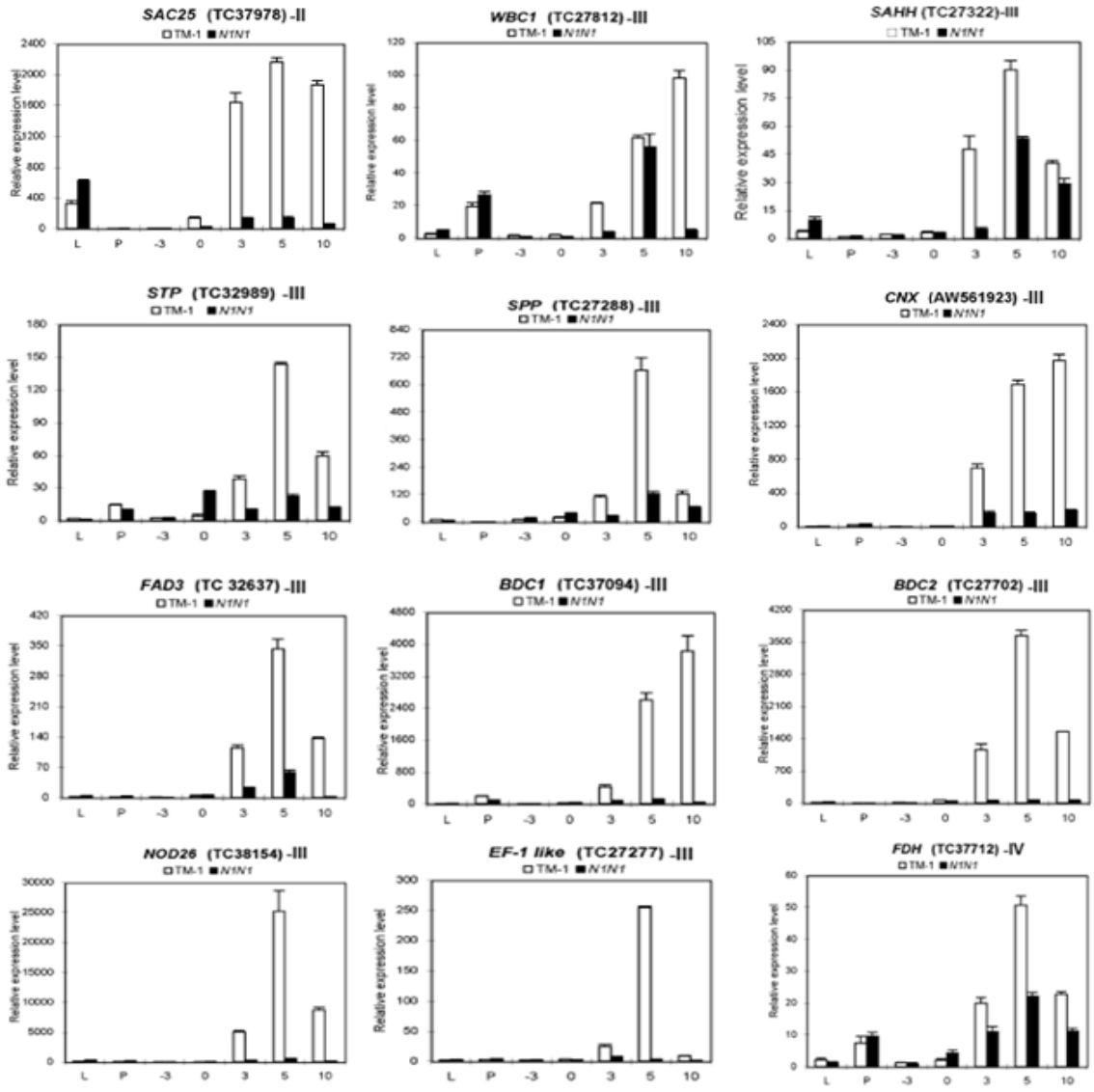

Figure 2.6. Histograms showing relative expression ratio of 17 fiber-associated genes detected by qRT-PCR. The expression patter (I-V) of each gene is indicated after gene name. The relative expression levels are displayed using the data shown Table 2. 6. L: leaves; P: petals; 0, 3, and 5: ovules at 0, 3, 5 DPA, respectively; 10: fibers at 10 DPA.

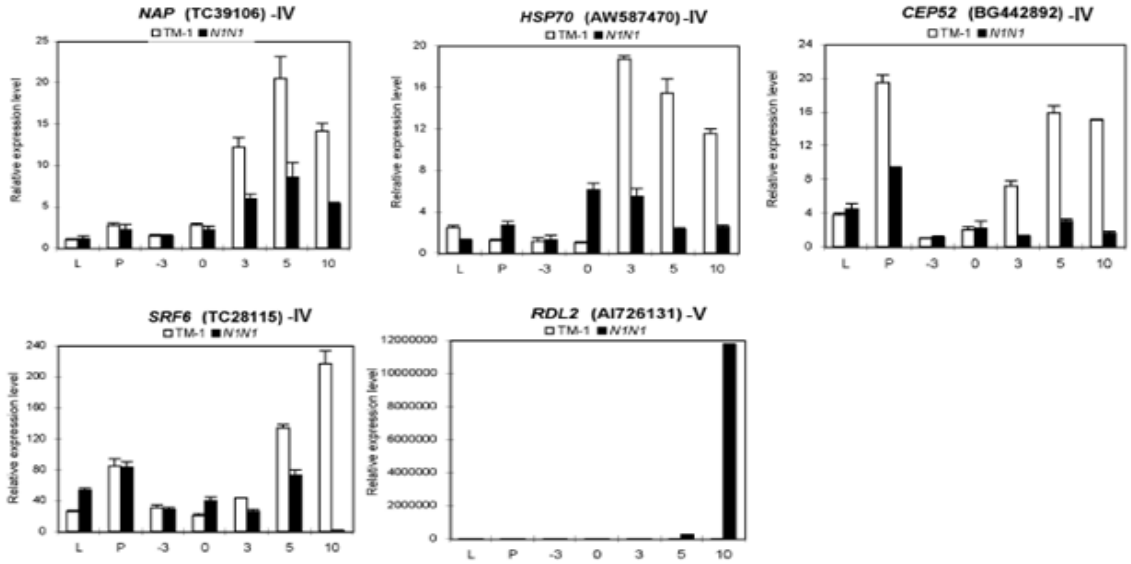

Figure 2.6. Histograms showing relative expression ratio of 17 fiber-associated genes detected by qRT-PCR. Continued.

Table 2.6. Quantitative RT-PCR analysis of relative expression levels of 23 fiber-associated genes

EST ID	Leaf		Petal		-3DPA		0DPA		3DPA		5DPA		10DPA	
	TM-1	*N1N1*	TM-1	*N1N1*	TM-1	*N1N1*	TM-1	*N1N1*	TM-1	*N1N1*	TM-1	*N1N1*	TM-1	*N1N1*
TC 27364	3.60± 0.67	7.01± 0.66	21.51± 5.05	17.79± 3.07	1.12± 0.02	1.00± 0.22	20.07± 1.40	4.10± 1.75	603.27± 64.46	292.71± 1.17	3304.00± 234.98	573.37± 8.25	3549.34± 51.06	1226.22± 173.07
TC 37712	2.16± 0.52	1.44± 0.10	7.43± 2.05	9.54± 1.25	1.27± 0.04	1.00± 0.23	1.98± 0.49	4.30± 0.91	19.88± 1.73	11.00± 1.57	50.56± 2.86	22.01± 1.21	22.68± 0.74	11.11± 0.87
TC 37267	1.00± 0.04	2.04± 0.10	65.50± 1.45	73.69± 13.32	1.05± 0.07	1.05± 0.01	9.87± 0.10	5.70± 0.43	239.41± 13.68	24.99± 5.55	704.28± 50.83	251.31± 39.97	1694.53± 86.21	43.51± 3.18
TC 27812	2.44± 0.42	5.02± 0.12	19.38± 2.11	26.60± 1.93	1.54± 0.37	1.01± 0.02	1.93± 0.27	1.00± 0.07	21.36± 0.47	3.98± 0.06	61.53± 1.60	56.10± 7.78	98.13± 4.42	4.76± 0.70
TC 27322	3.70± 0.66	9.87± 1.80	1.00± 0.04	1.29± 0.23	2.13± 0.07	1.53± 0.54	3.06± 0.66	2.82± 0.46	47.61± 7.08	4.81± 0.94	89.99± 4.94	53.32± 1.13	40.32± 1.39	29.24± 2.90
TC 32989	1.81± 0.01	1.00± 0.21	14.35± 0.63	9.08± 1.13	2.50± 0.02	2.62± 0.26	4.14± 1.35	26.48± 0.87	38.14± 2.78	10.39± 0.18	143.34± 1.52	21.86± 1.93	59.44± 3.78	11.69± 0.60
TC 39106	1.00± 0.11	1.11± 0.31	2.71± 0.28	2.19± 0.64	1.44± 0.18	1.45± 0.11	2.75± 0.19	2.18± 0.43	12.15± 1.18	5.94± 0.58	20.44± 2.68	8.53± 1.76	14.12± 0.93	5.33± 0.17
TC 27288	10.29± 1.02	8.30± 0.44	1.00± 0.03	1.16± 0.12	11.60± 0.20	18.42± 3.88	16.11± 5.66	38.05± 2.27	107.63± 9.30	27.16± 2.27	660.16± 55.24	123.35± 9.88	121.94± 13.81	63.41± 4.41
TC 38383	1.15± 0.10	1.59± 0.92	5.72± 0.22	14.52± 5.11	2.16± 0.13	1.00± 0.18	2562.49± 112.92	539.94± 46.40	2659.00± 74.75	785.06± 84.29	2067.02± 80.85	3236.01± 134.63	243.88± 7.31	1376.38± 91.68
AI 726131	1.00± 0.04	3.86± 0.53	99.73± 4.57	70.52± 3.07	1.52± 0.09	2.80± 0.03	25.40± 0.63	28.38± 7.31	241.63± 33.94	32.22± 7.09	556.41± 22.84	262750.38 ± 5464.23	1837.25± 108.96	11727021. 77 ± 46984.56
TC 37978	325.29± 38.10	623.67± 19.47	3.00± 0.04	4.02± 0.47	6.98± 0.05	7.40± 0.31	139.82± 11.02	26.65± 3.44	1638.80± 124.69	136.99± 5.16	2161.78± 59.89	134.21± 22.18	1864.90± 58.34	60.22± 2.01
AW 561923	7.75± 0.24	14.72± 1.36	35.67± 0.74	50.10± 2.48	5.02± 0.99	1.61± 0.46	13.18± 1.48	10.31± 0.85	1117.97± 78.62	261.38± 27.43	2721.15± 75.98	254.82± 25.68	3169.41± 122.97	309.40± 17.46
TC 27269	1.19± 0.24	2.66± 0.88	2.90± 0.21	3.55± 0.12	1.59± 0.93	2.01± 0.58	2.35± 0.95	2.40± 0.30	5.97± 0.83	2.28± 0.10	10.39± 1.01	1.52± 0.04	8.90± 1.84	1.58± 0.30
TC 39685	2.46± 0.20	1.26± 0.08	1.21± 0.14	2.68± 0.38	1.10± 0.38	1.27± 0.43	1.00± 0.07	6.11± 0.65	18.68± 0.34	5.46± 0.78	15.42± 1.36	2.30± 0.15	11.50± 1.36	2.51± 0.17
BG 442892	3.78± 0.20	4.45± 0.64	19.47± 0.91	9.38± 0.04	1.00± 0.03	1.06± 0.20	1.99± 0.40	2.16± 0.85	7.16± 0.64	1.23± 0.13	15.85± 0.88	3.01± 0.26	15.03± 0.10	1.61± 0.19
TC 32637	2.54± 0.16	4.31± 0.61	1.23± 0.18	2.85± 0.79	1.55± 0.17	1.00± 0.05	4.25± 1.01	5.31± 0.93	114.56± 7.16	22.78± 1.03	342.51± 22.80	55.35± 4.92	136.87± 2.91	2.35± 0.19
TC 28115	25.63± 1.17	53.57± 2.62	84.64± 9.63	83.29± 6.78	30.41± 3.70	28.44± 2.56	20.07± 2.93	39.49± 5.11	43.51± 0.00	25.46± 3.13	133.44± 4.18	73.01± 6.84	216.77± 16.58	1.00± 0.58
TC 27374	25.87± 3.15	22.01± 0.09	1.00± 0.04	3.91± 0.41	93.27± 7.75	197.18± 5.14	66.87± 6.83	86.62± 10.03	636.20± 15.98	241.63± 16.23	1192.69± 112.03	471.14± 30.42	651.07± 27.40	13.18± 1.26
TC 37094	2.22± 0.07	4.21± 0.29	189.14± 16.31	96.34± 2.69	1.56± 0.01	1.00± 0.39	12.55± 1.26	22.68± 2.09	419.73± 52.26	80.82± 3.16	2598.27± 187.41	117.51± 5.18	3830.54± 379.19	30.98± 0.33
TC 27702	10.80± 1.71	25.05± 0.46	1.00± 0.22	1.34± 0.06	12.38± 0.17	8.22± 1.31	71.84± 1.49	50.10± 1.57	1157.40± 129.91	59.16± 1.03	3632.30± 125.92	67.18± 1.17	1552.09± 0.00	61.11± 7.31
TC 38154	76.64± 2.00	357.88± 16.66	53.82± 0.77	156.50± 21.83	1.31± 0.05	1.00± 0.12	1.31± 0.16	24.65± 2.39	5066.12± 174.89	286.69± 16.43	25180.16± 3429.41	596.34± 19.69	8679.12± 455.93	130.69± 15.93
TC 37642	1.00± 0.00	1.70± 0.02	26.60± 1.92	45.89± 6.98	6.08± 0.21	4.28± 0.87	5.31± 0.87	10.22± 1.68	83.67± 4.65	15.93± 0.50	1629.26± 82.28	3.35± 0.56	146.03± 11.83	8.42± 1.31
TC 27277	2.30± 0.08	3.22± 0.21	3.13± 0.08	3.85± 0.23	2.14± 0.04	2.37± 0.43	3.30± 0.62	2.54± 0.32	24.48± 3.13	6.36± 1.74	254.23± 2.03	3.94± 0.02	8.75± 0.36	1.00± 0.56

Figure2.6. Using gene-specific primers (Table 2.4), transcripts of each gene were amplified in seven tissues in TM-1 and *N1N1*, and the data were analyzed using relative expression ratios with standard deviations (Figure 2.5a, Table 2.6). The expression patterns of 21 genes detected by microarray analysis were confirmed by qRT-PCR analysis. Although for some genes the relative expression ratios detected by qRT-PCR and the fold changes detected by microarray, were not the same. The expression ratios detected in the qRT- PCR assays appeared to be much higher than those detected in the microarray analysis, suggesting that qRT-PCR is a relatively sensitive method for gene expression analysis. Alternatively, 70-mer oligos may cross-hybridize to the homologous genes in the allopolyploid genomes. For two genes, *SRF6* and *CEP52*, the expression ratios detected by qRT-PCR matched fold-changes in other comparisons but did not match the increased fold changes in the ovules at 3 DPA detected by microarray analysis (Table 2.5, Table 2.6). One possibility is that the primers designed for the two genes may not correspond to the specific members of multiple gene families encoding the receptor protein (SRF6) and ribosomal protein (CEP52).

The data reveal at least five types of gene expression patterns that are associated with temporal regulation of early fiber development in TM-1 and *N1N1* (Figure 2.5a, Table 2.5, Figure 2.6). First, *GhPDF1* expression was highly induced in the immature ovules at −3 DPA, prior to the formation of fiber cell initials, and remained highly expressed until 5 DPA. Protodermal factor 1 (PDF1) is a protein involved in cell fate determination. In *Arabidopsis*, *PDF1* is exclusively expressed in the L1 layer of

vegetative and floral meristems, in organ primordia, and in protodermal cells during embryogenesis. *PDF1* expression is undetectable in the epidermis of mature organs (Abe et al., 1999, 2001). Notably, *GhPDF1* was highly expressed in immature ovules (−3 DPA) in the *N1N1* mutant as in TM-1, whereas its expression was undetectable in the fibers at 10 DPA in *N1N1*.

Second, genes are activated in the ovules at 0 DPA. *GhMYB25* represents small group of the large MYB transcription gene family (Arabidopsis Genome Initiative 2000). *GhMYB25* expression was up-regulated in the fiber-bearing ovules but not in the non-fiber tissues. Its expression was induced in the ovules at 0 DPA and the transcripts were accumulated at high levels until 5 DPA. The mRNA levels declined when the fiber cells proceeded to the elongation stage (10 DPA). The accumulation of *GhMYB25* transcripts was highest in the young ovules (0 and 3 DPA), low in the fibers (10 DPA), and undetectable in the leaves, petals, and immature ovules (−3 DPA) (Fig 2.5). *GhMYB25* expression levels were decreased during early stages of fiber elongation (10 DPA). On the contrary, *GhMYB25* in the *N1N1* mutant lines transcripts were accumulated in the ovules from 0 to 5 DPA and remained at a relatively high level in fibers (10 DPA). A similar activation pattern was observed for *SAC25* that were up-regulated in the ovules at 0 DPA and remained highly expressed from 3 to 10 DPA.

Third, 12 of 23 genes were up-regulated starting from 3 DPA and their expression continued to rise and peaked at 10 DPA. These genes encode a variety of important

proteins such as fiber E6 protein, dehydration-induced protein, ABC transporter, S-adenosylhomocysteine hydrolase, and BURP domain-containing protein. *E6* encoding fiber protein E6 is one of the highly expressed genes in fiber tissues throughout the fiber development (John and Crow, 1992; John, 1996). E6 mRNA and protein are highly accumulated during stages of late primary cell wall and early secondary cell wall biosynthesis. *E6* expression was undetectable in leaves, petals, and immature ovules (0 DPA) but up-regulated 200-fold in the ovules at 3 DPA, continued to rise 900-fold in the ovules at 5 DPA, and peaked in the fibers at 10 DPA (Figure 2.5a, Table 2.6), confirming the gene encoding fiber protein E6 was expressed throughout the development of fiber cells (John and Crow, 1992). The fold induction of *E6* expression in the fiber-bearing ovules (3–5 DPA) and fibers (10 DPA) in *N1N1* was 1/5–1/3 of those in TM-1. Similarly, *RDL1* was up-regulated in the fiber-bearing ovules at 3 and 5 DPA and in the fibers at 10 DPA, whereas the expression levels were substantially lower in *N1N1* and non-fiber tissues (Figure 2.5a). *RDL1* promoter containing an L1 box and a MYB-binding motif is a fiber-specific gene (Li et al., 2002a; Wang et al., 2004). It is notable that the promoter of *RDL1* contains MYB-binding elements, indicating physical interactions between MYB transcription factors and *RDL1* (Wang et al., 2004). Collectively, the data support a model (Figure 2.5b) of sequential activation of many genes involved in various biological pathways, leading to the progression of fiber cell development.

Fourth, six genes were expressed in non-fiber tissues, but their expression levels were dramatically increased in the ovules and fibers. These genes encode proteins such as nucleosome assembly protein, elongation factor (EF-1), HSP70, receptor family, and ribosomal protein CEP52.

Fifth, up-regulation was found in only one tissue type. *ACT* and *RDL2* were highly expressed in TM-1 ovules at 5 DPA and *N1N1* fibers at 10 DPA, respectively. Significantly, the expression levels of 23 genes, except *RDL*, in seven tissues tested were much lower in *N1N1* than in TM-1. This suggests that the *N1N1* mutation disrupts the temporal regulation of many genes involved in various biological pathways including signals for fiber cell elongation and the number of cells committed to fiber differentiation.

DISCUSSION

Cotton fiber ESTs and their expression patterns in the fibers and non-fiber tissues

The majority of cotton ESTs are derived from the fibers at 6–10 DPA in *G. arboreum* L. or *G. hirsutum* L. species. Estimates indicate that fiber transcriptome represents 35–40% of the genes in the cotton genome (Arpat et al., 2004). It is unclear about the proportion of the cotton fiber ESTs that are expressed during the development of fibers and non-fiber tissues. Microarray analysis of gene expression using fibers and non-fiber tissues suggests that approximately 40–50% of the current fiber ESTs are expressed in leaves and flower petals (Figure2.4, Table 2.3). However, the number of genes expressed in

non-fiber tissues could be lower (≅20%) than 40% because many non-fiber tissues (e.g., roots and stems) were not used in the study. Among 51 genes that are differentially expressed between TM-1 and *N1N1* ovules (3 DPA), 28 are expressed either in leaves or in petals (Table 2.3, Figure 2.4a). Moreover, five of nine randomly selected genes that display expression differences between TM-1 and *N1N1* are also expressed in non-fiber tissues (leaves and petals) in qRT-PCR analysis (data not shown). The high percentage of ESTs that are expressed in non-fiber tissues indicates that many genes involved in general biological pathways, such as metabolism, energy production, and the biosynthesis of primary and secondary cell walls, are expressed throughout the development of vegetative tissues and reproductive organs. However, during fiber cell development their expression may be dramatically induced in response to rapid cell expansion and growth.

Consistent with the above notion, in a previous study using 12,227 fiber ESTs from *G. arboreum* L., Arpat et al. (2004) found that only 81 genes were up-regulated and 2,553 "expansion-associated" genes were down-regulated during the developmental switch from primary to secondary cell wall biosynthesis. This low percentage of up-regulated genes detected in the stages of fiber cell expansion and elongation is reminiscent of ≅100 genes detected in the early stages of fiber development (Figure 3.4, Table 2.3) regardless of the small set of genes used in this study. The data obtained from two experiments using different stages of fiber development are difficult to compare in order to detect a common set of genes.

Developmental regulation of fiber-associated genes in TM-1 and *N1N1* mutant

Fiber cell expansion and elongation occur continuously through a diffuse-growth mechanism (Tiwari and Wilkins, 1995). It is notable that the timing of up-regulation of 23 fiber-associated genes coincides with temporal control of fiber cell development, ranging from −3 DPA (1 gene), 0 DPA (2 genes), 3 DPA (12 genes), to 5 or 10 DPA (2 genes) (Figure 2.5a, Figure 2.6). Six of the 23 genes were up-regulated in the ovules at 3 DPA and in the fibers at 10 DPA but also expressed in the non-fiber tissues such as leaves and petals. Thus, ovular development at 3 DPA is a critical step for rapid cell expansion and cellular growth. The data support a model of temporal activation of regulatory networks during early stages of fiber cell development (Figure 2.5b). External and internal signals for fiber cell differentiation may be transmitted to fiber cell primordia, which leads to the activation of "patterning" genes (Hülskamp, 2004) including *GhPDF1*. The *Arabidopsis* homolog of *GhPDF1* encoding a putative extracellular proline-rich protein is exclusively expressed in the L1 layer of shoot apices and the protoderm of organ primordia (Abe et al., 1999, 2001). Molecular events of gene activation in fiber cell primordia may be coupled with sequential activation of transcription factors and proteins such as MYB transcription factors and RDL1 proteins that are important to fiber or trichome cell differentiation. There is evidence that cotton and *Arabidopsis* use similar transcription factors for regulating trichome development. Notably, GaMYB2 and RDL1 promoter interact physically and overexpressing *GaMYB2* complements the *gl1* mutant

phenotype as well as induces the development of seed trichomes in *Arabidopsis* (Wang et al., 2004). The data suggest a critical role of MYB transcription factors in fiber cell differentiation. Interestingly, *RDL1* and *MYB25*, a homolog of *AtMYB17* and *106* that are in a separate branch but closely related to *GL1*, were up-regulated in the fiber-developing ovules at 0 and 3 DPA, respectively. This suggests that multiple components of MYB transcription factors and other proteins are involved in fiber cell differentiation. Up-regulation of many other genes such as *WBC1*, *FDH*, *EF1A*, and *NOD26* (Table 2.5, Table 2.6) is likely involved in late stages of fiber cell differentiation. For example, *GhWBC1* encoding an ATP-binding transporter of the WBC subfamily is expressed at low levels in the ligon-lintless mutant that is defective in fiber cell elongation (Zhu et al., 2003).

N1N1 mutation not only delays the onset of fiber cell initiation by 12–24 h but also reduces the number of undifferentiated protodermal cells that develop fiber cell initials. We identified 117 genes and 30 genes that displayed expression changes in the *N1N1* mutant at 3 DPA and 0 DPA (Table 2.3). When per-gene variance is used, 498 and 195 genes were significantly different between the mutant and wild type in these two stages. Moreover, equal number of genes were up- or down-regulated in the ovules at 3 DPA, whereas 20 genes were up-regulated and 10 down-regulated in the ovules at 0 DPA in *N1N1*. Relatively equal number of genes that were up- or down-regulated in *N1N1* (Table 2.3) suggest that the *N1N1* mutation has both positive and negative effects

on gene regulation in the early stages of fiber development, leading to fewer fibers and a defective process of fiber cell elongation.

Twenty-three genes, with a few exceptions, were expressed exclusively in fiber-related tissues including ovules (–3 and 0 DPA), fiber-bearing ovules (3 and 5 DPA), and fibers (10 DPA) and down-regulated in the *N1N1* mutant, which are the likely candidate genes involved in the process of fiber cell formation. All 23 fiber-associated fiber genes were up-regulated in the ovules and fibers compared to the non-fiber tissues in TM-1. Compared to TM-1, the expression levels of these genes, except for *RDL2* and *GhMYB25*, were dramatically reduced in the corresponding tissues tested in *N1N1* (Figure 2.5a, Table 2.6), suggesting that temporal regulation of gene activation is disrupted in the *N1N1* mutant.

We note that for a technical difficulty in dissecting fibers from *N1N1* mutants, we used fiber-bearing ovules at 10 DPA in the mutant compared to fibers at 10 DPA in the wild type, which may obscure the detection of fiber-related genes. For example, *RDL2* was expressed only in the fibers at 10 DPA in *N1N1* but not in TM-1. *GhMYB25* was expressed at higher levels in the fiber-bearing ovules at 5 DPA and fibers at 10 DPA in *N1N1* than in TM-1. The expression differences detected at 10 DPA may be caused by different tissues used in the study (fibers in TM-1 and fiber-bearing ovules in *N1N1*). Alternatively, *GhMYB25* and *RDL2* expression may be affected by negative regulators or affected indirectly by interacting with other protein factors induced by the *N1N1* mutation,

which is reminiscent of the equal number of the genes that were up- or down-regulated in the *N1N1* mutant (Table 2.3).

FDH encodes an enzyme involved in the synthesis of long-chain lipids found in the cuticle (Lolle et al., 1992; Pruitt et al., 2000). The mutations in *FDH* suppress epidermal cell interactions in *Arabidopsis*, exhibiting a deleterious effect on trichome development (Yephremov et al., 1999). The *FDH*-like gene is highly expressed in developing fibers (Li et al., 2002a) and is repressed in the *N1N1* mutant, suggesting down-regulation of the *FDH*-like gene in the *N1N1* background might be associated with abnormal development of fiber cell initials. Down-regulation of the genes (Figure 2.5a, Figure 2.6) involved in cell differentiation (e.g., *PDF1*), transcriptional regulation (e.g., *MYB25*), signal transduction (e.g., *SPP*,*BDC1*, and *BDC2*), transport facilitation (e.g., *STP* and *WBC1*), cell wall biosynthesis (e.g., *FDH*,*RDL1*, and *NOD26*), and translational regulation (e.g., *EF1A* and *EF1-like*) are likely to be associated with a series of defective processes of fiber cell formation and elongation in the *N1N1* mutant. *ACT* transcripts were highly accumulated in TM-1 ovules at 5 DPA, suggesting a role of actin cytoskeleton during fiber development (Li et al., 2005). Notably, S-adenosylhomocysteine hydrolase is a key enzyme involved in the intracellular methylation reactions (Tanaka et al., 1997; Fojtova et al., 1998). Suppression of *SAHH* may lead to pleiotropic effects on polarized growth of developing fiber cells in the *N1N1* mutant. Our data suggest that the *N1* mutation affects many downstream genes involved in various biological pathways including cell differentiation, transcriptional and

translational regulation, and signal transduction that are essential for the determination of the number of fiber cells and fiber cell elongation during early stages of fiber development.

METHODS

Plant materials

G. hirsutum L. cv. TM-1 and its isogenic *N1N1* mutant lines (>S6 generation) were grown in a greenhouse. Young emerging leaves (about 2 inches in diameter) were collected from seedling plants, and petals were collected at the day of anthesis (0 DPA) before the flower color changed from white to pink. Flower buds prior to anthesis (−3 DPA) were collected when the ovules were enclosed by squares of 1/3–1/2 inches in diameter. Flowers were tagged on the day of anthesis, and ovules were harvested at 0, 3, and 5 DPA. Cotton bolls (TM-1) at 10 DPA were harvested for dissecting fibers, whereas for *N1N1* the developing ovules were harvested. For each genotype, I used two biological pools, each with ten plants grown at similar stages. Leaves were harvested by pooling one leaf from each of ten individual plants. Ovules or fibers were dissected from five bolls collected in each of ten plants. The fresh tissues were frozen in liquid nitrogen and stored in a −70°C freezer or subjected to RNA extraction.

Scanning Electron Microscopy (SEM)

SEM was performed using a modified protocol (Murai et al., 2002; Tian et al., 2003). In brief, ovules from 0 to 3 DPA were dissected from immature ovules of TM-1 and the isogenic naked seed mutant, *N1N1*. The ovules were fixed in a solution containing 3% each of formaldehyde and glutaraldehyde in 0.1 M sodium cacodylate buffer (pH = 7.4) and rinsed in 0.2 M sodium cacodylate buffer (pH = 7.4) three times. The ovules were washed in an ethanol series from 10 to 70% every 15 min. The concentration of the ethanol was increased to 100% within 18 h to dehydrate the samples for SEM analysis. The specimens were prepared by critical-point, dried with CO_2 at 1,400 and 1,800 psi, consecutively, and mounted by conductive gold paint and sputter coating. The samples were then scanned and analyzed using a JEOL (JSM-6400) SEM located in the Microscopy and Imaging Center at Texas A&M University, with an accelerating voltage of 15 kV and a working distance of 39 mm. Images were scanned and stored as TIFF files.

EST sequence analysis and oligo selection

To construct a pilot set of cotton 70-mer oligos, I selected a subset of genes based on *Arabidopsis* protein sequences (Arabidopsis Genome Initiative 2000) using TIGR Gene Indices (http://www.tigr.org/tdb/tgi/plant.shtml). Each family of the *Arabidopsis* proteins was analyzed using BLAST against the entire cotton EST database with a method similar to the previously published (Blanc and Wolfe, 2004a, 2004b). Two proteins were considered to be similar if their amino-acid sequences shared over 38% identity or 60%

similarity when the aligned sequences were ≤ 100 amino acids (a.a.) in length, or over 28% identity or 50% similarity when the aligned sequence were > 100 a.a. in length (*E*-value ≤0.00001). The selected ESTs were compared using BLAST against each other to eliminate duplicates that shared more than 95% sequence identity (Table 2.1). The remaining ESTs were used for 70-mer oligo design so that each oligo had a minimum of secondary hairpin structures (Kane et al., 2000; Rouillard et al., 2002). Sequence lengths of 70 nucleotides with a similar melting temperature of 73±2°C were selected within 1,000 nucleotides of the 3′ end of predicted coding sequences by using the software developed by ProbeSelect (Li and Stormo, 2001) or Featurama (http:// probepicker.sourceforge.net/). We selected 1,334 oligos from cotton ESTs (http://www. tigr.org/tdb/tgi/plant.shtml), 15 controls, and 121 and 66 oligos designed from annotated *Arabidopsis* chloroplast and mitochondrial genes, respectively, using the Organelle Genome Resources at the NCBI http://www.ncbi.nlm.nih.gov/genomes/ORGANELLES/ organelles.html. A total of 1,536 features were spotted in duplicate on each slide.

Microarray experimental design and statistical analysis

The 70-mer oligos were synthesized at Operon (http://www.operon.com/arrays /omad.php). The oligo solution (30 μM in 3XSSC) was spotted onto SuperAmine microarray slides (SMM) using micro spotting Stealth 3 pins (SMP3) (TeleChem International, Inc. Sunnyvale, CA) in an OmniGrid Accent microarrayer (GeneMachines,

San Carlos, CA). Microarray design and statistical tools were developed as previously described (Lee et al., 2004; Wang et al., 2005a). Messenger RNA (mRNA) was directly isolated from various tissues of TM-1 and *N1N1* using Poly(A) Pure (Ambion, Austin, TX) according to the manufacturer's recommendations (Wang et al., 2005a). Cy3- and Cy5-labeled probes were generated using the CyScribe Post-labelling kit (Amersham Biosciences, Piscataway, NJ). Messenger RNA was reverse-transcribed into amino allyl-labelled cDNA. The single-strand cDNAs were coupled with CyDye-NHS ester, which binds to the modified nucleotides. We used 500 ng of mRNA in each labeling reaction using Cy3- or Cy5-dCTP (Amersham Biosciences, Piscataway, NJ). We used dye-swap experimental design and linear model (Kerr and Churchill, 2001; Chen et al., 2004) for our microarray analysis. For one dye-swap experiment, we used two sets or four labeling reactions. After labeling, one Cy3-dCTP reaction is mixed with one Cy5-dCTP reaction to make one probe. Therefore, two "identical" probes, each containing an equal amount of Cy3- and Cy5-labled cDNAs, were hybridized with two slides, which constituted one dye-swap experiment. The dye-swap was repeated once as a technical replication. The large dye-swap (four slides) was repeated using another biological sample (e.g., RNAs isolated from different pools of 3-DPA ovules). Therefore, each experiment consists of four technical replications and two biological replications in a total of eight slides (Chen et al., 2004). Hybridization was performed overnight (◎14 h) at 65°C. After hybridization, the slides were washed twice for 4 min each in 2X SSC, 0.2% SDS, again twice for 2 min each in 0.2XSSC, and twice for 2 min each in 0.05X SSC. After drying the slides by brief

centrifugation (5 min at 850 rpm), the slides were scanned using GenePix 4000B (Axon, Foster City, CA), and the images were captured by GenePix Pro 4.1 software.

After the data were processed using natural logarithm ratios of green and red hybridization signals, a robust and locally weighted linear regression (lowess) (Cleveland, 1979) was used to remove non-linear components (e.g. dye and pin effects) (Quackenbush, 2002). For the duplicate spots in each feature, we used an average value for data analysis. No additional steps for data normalization and background subtraction are needed for the the analysis of variance (ANOVA) model (Lee et al., 2004). The data were then subjected to AVONA test in a linear model to estimate the significant changes in gene expression caused by the two treatments (genotypes) (Black and Doerge, 2002; Lee et al., 2004). A standard *t*-test statistic was used for this comparison based on the normality assumption for the residuals. The standard false discovery rate (FDR) (Hochberg and Tamhane, 1987) was applied to control multiple testing errors using a significance level $\alpha = 0.05$.

Quantitative RT-PCR (qRT-PCR) analysis

Messenger RNA (mRNA) was extracted from seven different tissues including leaves, petal, ovules at −3 DPA, 0 DPA, 3 DPA, 5 DPA, and fibers at 10 DPA in TM-1 and the *N1N1* mutant using Poly(A) Pure (Ambion, Austin, TX). For microarray data verification, the same mRNA used in the microarray analysis was used for qRT-PCR analysis. The

cDNAs were amplified by Superscript II reverse transcriptase reaction (Invitrogen, Carlsbad, CA). For the transcript amplification, gene-specific primers (Table 2.4) were designed using Primer Express version 2.0 software. The qRT-PCR reaction was carried out in a final volume of 20 µl containing 10 µl SYBR Green PCR master mix (Applied Biosystems, Foster City, CA), 1 µM forward and reverse primers, and 0.1 µM cDNA probe in a ABI7500 Real-Time PCR system (Applied Biosystems, Foster City, CA). Cotton *HISTONE3* (AF024716) was used to normalize the amount of gene-specific RT-PCR products (Wang et al., 2004). All reactions were performed in three replications. The amplification data were analyzed using ABI7500 SDS software (version 1.2.2), and the fold changes were calculated using the standard in each reaction.

CHAPTER III

TRANSCRIPTOME AND FUNCTIONAL ANLAYSES OF FIBER GENES IN COTTON

INTRODUCTION

Cotton is the major source of natural fibers. Cotton fiber development undergoes four distinct and overlapping processes: fiber cell initiation, elongation, cellulose biosynthesis, and maturation (Basra and Malik, 1984; Tiwari and Wilkins, 1995; Wilkins and Jernstedt, 1999). Although fiber initiation is a fundamental step for cell differentiation and fiber development, little is known about the molecular basis of fiber cell initiation (Lee et al., 2006; Taliercio and Boykin, 2007).

Fiber cell differentiation begins prior to anthesis (-2 DPA) and primordial fiber cells destined to become lint fibers emerge at the anthesis (Lee et al., 2006; Taliercio and Boykin, 2007). The fiber initials (or primordial cells in the epidermis) during fiber initiation are only a small portion of the ovule cells which makes technical difficulties during tissue preparation. To overcome this limitation, laser capture microdissection (LCM) can be used. Laser capture microdissection allows the harvest of individual cells from surrounding cells and the application of LCM for tissue-specific analysis has been

reported in various studies (Matsunaga et al., 1999; Kerk et al., 2003; Nakazono et al., 2003; Rook et al., 2004; Casson et al., 2005; Angeles et al., 2006; Cai and Lashbrook, 2006). In cotton, the fiber initials were collected from inner ovules at 0 DPA using LCM (Wu et al., 2006). This study found that *GhMYB25* and *GhHD1* transcripts were abundant in the fiber initials compared to the epidermal cells on the day of anthesis (Wu et al., 2006). Coupled with aRNA amplification and microarray analysis, LCM was successfully applied in the analysis of global gene expression in specific cells (Jenson et al., 2003; Nakazono et al., 2003; Park et al., 2004; Casson et al., 2005; Cai and Lashbrook, 2006; Zhu et al., 2006; Day et al., 2007).

The *Arabidopsis* trichome is an excellent model system to study cotton fiber initiation. Cotton fibers are seed trichomes derived from the ovule epidermis and seed trchome development in cotton shares many similarities with leaf trichome development in *Arabidopsis* where trichome development has been extensively studied (Lee et al., 2007). Trichome development in *Arabidopsis* leaves is initiated by a 'trichome promoting complex' that consists of *GLABROUS1* (*GL1*), *TRANSPARENT TESTA GLABRA1* (*TTG1*) and *GLABRA 3* (*GL3)* (Szymanski et al., 2000). Trichome development in the shoot or leaf epidermis requires *GL2* expression in the cells (Szymanski et al., 2000; Hülskamp, 2004). The close relationship between cotton fibers and *Arabidopsis* trichomes was shown in the study of a cotton fiber MYB gene *GaMYB2*, a putative MYB transcription factor in diploid cotton species (*G. arboretum*) (Wang et al., 2004). *GaMYB2* expression driven by the *GL1* promoter in the *gl1* mutant complemented its

glabrous phenotype. Furthermore, *35S::GaMYB2* induced seed trichome production (Wang et al., 2004), suggesting cotton and *Arabidopsis* use similar transcription factors for trichome development. Two putative homologues of *Arabidopsis* trichome-associated genes, *TRANSPARENT TESTA GLABRA1* (*TTG1*), *GhTTG1* and *GhTTG3*, rescued the glabrous *ttg-1* phenotype in *Arabidopsis*, implying that similar molecular mechanisms operate in *Arabidopsis* trichome development and in cotton fiber development (Humphries et al., 2005).

In the previous chapter, possible "fiber-associated" genes have been identified using a pilot set of microarray (Lee et al., 2006). It provided a general overview of gene regulation in the wild type cotton and the fiberless *N1N1* mutant. However, a large scale global gene expression study using fiber-specific tissues is needed because the pilot set only covered 1,334 genes and these were mainly derived from 5-7 DPA ovules.

In this chapter, a larger array which includes a total of 22,789 oligos was used for gene expression analysis in the fiber cells during fiber initiation. Comparisons of laser capture microdissected tissues among different developmental stages uncovered several sets of genes including the *Gossypium hirsutum RD22-like* gene, *GhRDL*. In *Arabidopsis*, overexpression of a cotton fiber gene, *GhRDL,* was able to induce seed hair developement as well as an increased number of leaf trichomes. This result shows that *GhRDL* can affect seed hair development in *Arabidopsis* and implies that it may be a key player in cotton fiber development.

RESULTS

Laser capture microdissection for isolating fiber initials from ovules

Investigating genes that are highly expressed during fiber initiation has been challenging because of the specificity of gene expression in the biological materials. Several dozens of "fiber-associated" genes have been identified in the previous study (Lee et al., 2006). However, more direct evidence is needed to call these as "fiber-associated" genes because the tissues used in the microarray experiments were whole ovules, which include epidermal layers as well as inner ovules. To resolve this problem, it is desirable to use laser capture microdissection (LCM) to separate epidermal layers bearing fiber initials from inner ovule cells during the early event of fiber initiation. LCM provides a rapid means of isolating large homogeneous populations from heterogeneous tissues through direct visualization of the cells (Kerk et al., 2003; Nakazono et al., 2003; Espina et al., 2006). For the comparisons with fiber initials (or epidermis) at -2 DPA, 0 DPA and 2 DPA, LCM conditions were optimized for the epidermis and inner ovules of cotton and about 1,000 cells were captured (Figure 3.1). Antisense RNA amplification was conducted with captured cells because microarray analysis requires microgram amounts of a probe that uniformly represents the RNA population. It has been demonstrated that aRNA from LCM makes it feasible to perform global gene expression analyses using microarray (Nakazono et al., 2003; Park et al., 2004).

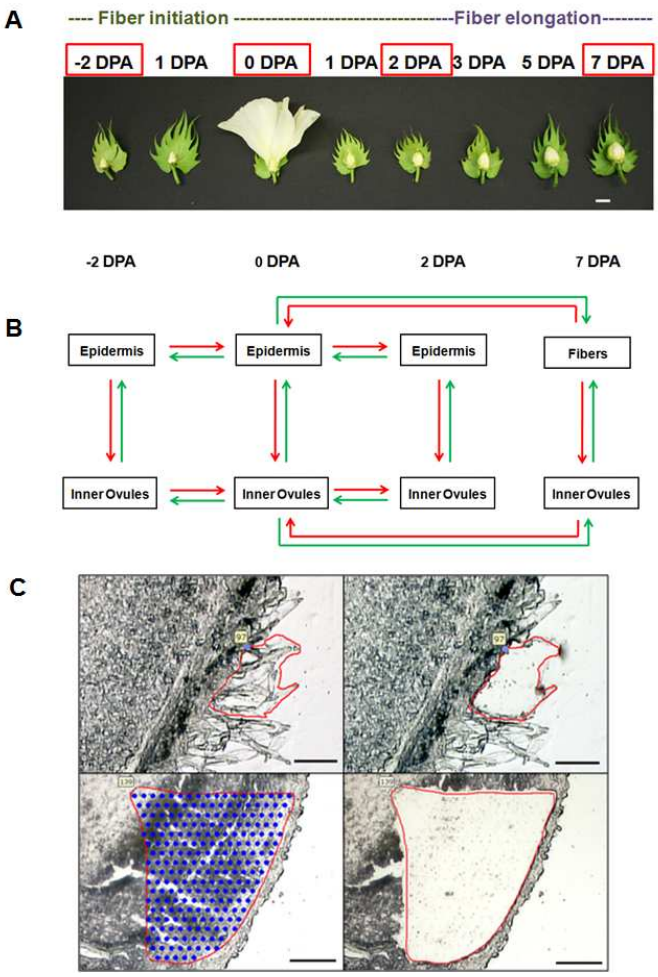

Figure 3.1. Microarray experimental design and application of laser capture microdissection

(A) Development of cotton fiber.

(B) Microarray experimental design. Arrows denote dyes that were used in each of the hybridization (red: Cy5, green: Cy3).

(C) Application of laser capture microdissection (LCM).

Differentially expressed genes in the epidermal layers of TM-1

Several hundred genes were detected in each microarray experiment (Table 3.1). In the comparison of epidermis vs. inner ovules (experiment 1-4 in Table 3.1), the number of differentially regulated genes during fiber initiation (-2 DPA, 0 DPA, and 2 DPA) was higher than that in fiber elongation (7 DPA). This may imply that this large set of microarray mainly designed from an ovule EST library is a good representation of the genes involved in fiber initiation. Differentially regulated genes in each stage (-2 DPA, 0 DPA, and 2 DPA) overlapped each other (Figure 3.2). Wilcoxon rank-sum test showed that the lists of differentially expressed genes at -2 DPA and 0 DPA are statistically similar. The lists at 0 DPA and 2 DPA are also similar whereas those at -2 DPA and 2 DPA are not (Table 3.2). In the epidermis comparisons, the differentially regulated genes at each developmental stage (-2 DPA, 0 DPA and 2 DPA) are similar each other (Table 3.2).

The microarray experiment uncovers several sets of interesting genes which may be involved in fiber initiation (Table 3.3). Among the 30 most abundant transcripts in GH_TMO library (Yang et al., 2006), 14 of which (approximately 46%) are found to be differentially regulated in inner ovule versus epidermis comparison. These include genes encoding protodermal factor 1, peroxidase precursor, alpha-expansin precursor, alpha-

Table 3.1. Number of differentially expressed genes

Experiment	Comparison	No. of up-regulated genes[a]	No. of down-regulated genes[a]
1	-2 DPA (E) vs. -2 DPA (O)	622	430
2	0 DPA (E) vs. 0 DPA (O)	972	742
3	2 DPA (E) vs. 2 DPA (O)	668	268
4	7 DPA (F) vs. 7 DPA (O)	432	428
5	0 DPA (E) vs. -2 DPA (E)	371	237
6	0 DPA (E) vs. 2 DPA (E)	53	412
7	0 DPA (E) vs. 7 DPA (F)	1207	710
8	0 DPA (O) vs. -2 DPA (O)	794	195
9	0 DPA (O) vs. 2 DPA (O)	104	293
10	0 DPA (O) vs. 7 DPA (O)	255	603

E: epidermis, O: inner ovules, F: fibers; [a]Up-regulation and down-regulation are based on the expression of epidermis (for experiment 1-4) or the expression of 0 DPA tissues (for experiment 5-10).

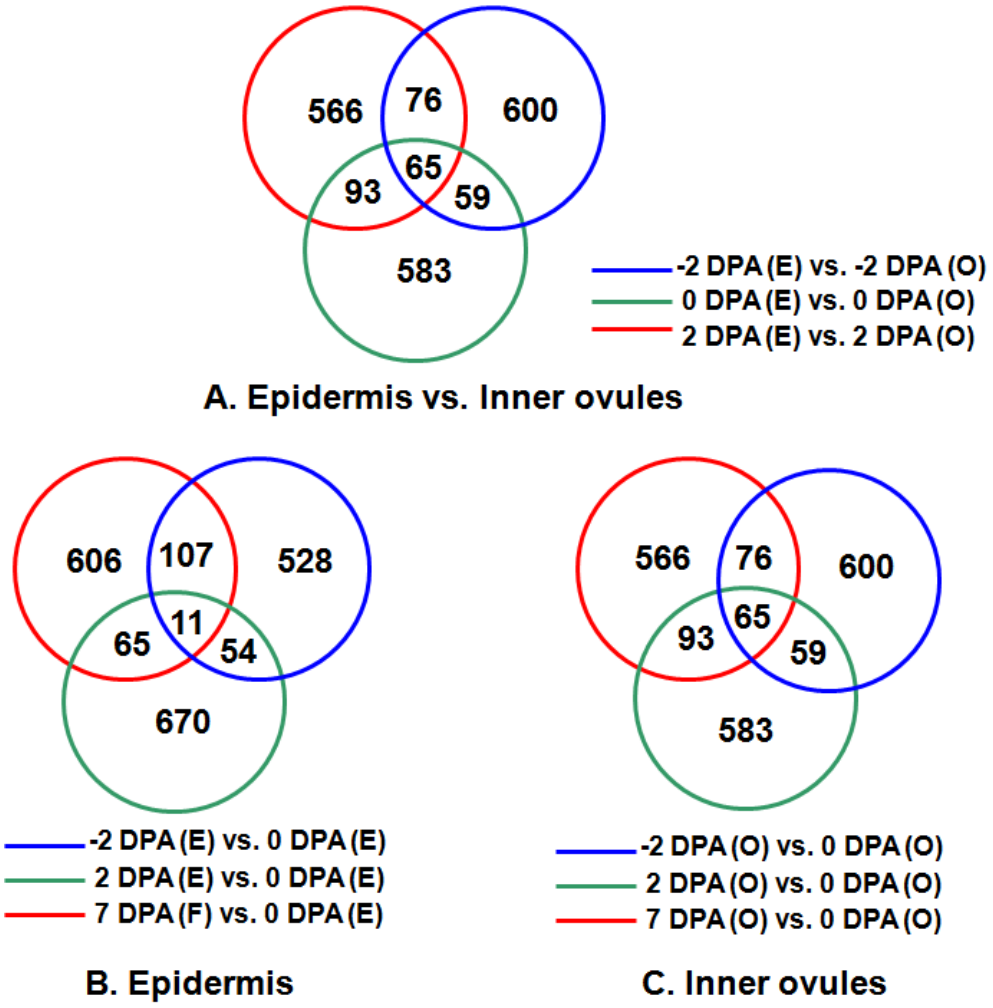

Figure 3.2. Comparative analysis of differentially expressed genes detected by microarray analysis.

Number of differentially regulated genes in the epidermis compared to the inner ovules at -2 DPA (blue), 0 DPA (green), and 2 DPA (red) **(A)**, in the epidermis at -2 DPA (blue), 2 DPA (green), and 7 DPA (red) compared to the epidermis at 0DPA **(B)**, in the inner ovules at -2 DPA (blue), 2 DPA (green), and 7 DPA (red) compared to the inner ovules at 0 DPA. The proportion shown in the diagrams may not be scaled.

Table 3.2. Wilcoxon rank sum test analysis[a] of differentially expressed genes

Comparison	Test	Hypothesis	P value	Conclusion[c]
Epidermis vs. inner ovules	1	Ho: μ1[b](-2 DPA) = μ2 (0 DPA) Ha: μ1 (-2 DPA) ≠ μ2 (0 DPA)	0.002	μ1 and μ2 are similar.
	2	Ho: μ2 (0 DPA) = μ3 (2 DPA) Ha: μ2 (0 DPA) ≠ μ3 (2 DPA)	0.701	μ2 and μ3 are similar.
	3	Ho: μ3 (2 DPA) = μ1 (-2 DPA) Ha: μ3 (2 DPA) ≠ μ1 (-2 DPA)	0.000	μ3 and μ1 are different.
Epidermis	1	Ho: μ5 (-2 DPA) = μ6 (2 DPA) Ha: μ5 (-2 DPA) ≠ μ6 (2 DPA)	0.676	μ5 and μ6 are similar.
	2	Ho: μ6 (2 DPA) = μ7 (7 DPA) Ha: μ6 (2 DPA) ≠ μ7 (7 DPA)	0.742	μ6 and μ7 are similar.
	3	Ho: μ7 (-2 DPA) = μ5 (2 DPA) Ha: μ7 (-2 DPA) ≠ μ5 (2 DPA)	0.895	μ7 and μ5 are similar.
Inner ovules	1	Ho: μ8 (-2 DPA) = μ9 (2 DPA) Ha: μ8 (-2 DPA) ≠ μ9 (2 DPA)	0.040	μ8 and μ9 are similar.
	2	Ho: μ9 (-2 DPA) = μ10 (2 DPA) Ha: μ9 (-2 DPA) ≠ μ10 (2 DPA)	0.000	μ9 and μ10 are different.
	3	Ho: μ10 (-2 DPA) = μ8 (2 DPA) Ha: μ10 (-2 DPA) ≠ μ8 (2 DPA)	0.015	μ10 and μ8 are similar.

[a]Tests were performed with differentially expressed genes (400 of up-regulated genes and down-regulated genes, total 800 genes) in each experiment.

[b]The number indicates the experiment number in Table 3.1.

[c]The conclusions were drawn based on the p value of 0.001.

Table 3.3. A subset of differentially expressed genes in the epidermis or fibers

Developmental stages	EST ID	Tentative annotation	FC
-2 DPA	TC77549	fiber protein Fb34 (*Gossypium barbadense*)	1.62
	TC73204	sucrose synthase	1.53
	TC60481	cellulose synthase	1.79
	TC80190	AtMYB3	1.43
	NP385922	AtMYB12	1.42
	BF276432	MYB transcription factor	1.33
	TC60660	MYB transcription factor	2.52
	TC62161	MYB transcription factor	2.25
	DT568260	AtARF11 (AUXIN RESPONSE FACTOR 11)	1.53
	TC66508	Auxin/aluminum-responsive protein	2.37
	TC65885	Ethylene-responsive protein	1.75
	TC63148	RTE1 (REVERSION TO ETHYLENE SENSITIVITY1)	2.02
	TC77366	abscisic acid-responsive HVA22 family protein	2.75
	TC77968	BRI1 (BRASSINOSTEROID INSENSITIVE 1)	1.37
0 DPA	TC75692	fiber protein Fb27 (*Gossypium barbadense*)	2.04
	DR457077	fiber protein Fb34 (*Gossypium barbadense*)	1.82
	TC62647	fiber protein Fb34 (*Gossypium barbadense*)	1.51
	TC73227	sucrose synthase	2.03
	TC66213	cellulose synthase	1.32
	TC70562	CSLC12 (CELLULOSE SYNTHASE-LIKE C12)	1.44
	AW667548	AtMYB60	1.32
	TC62952	GhMYB36	1.61
	TC61557	GhMyb38	1.33
	TC76794	ARF2 (AUXIN RESPONSE FACTOR 2)	1.30
	TC63397	ARF6 (AUXIN RESPONSE FACTOR 6)	1.31
	TC65581	auxin-responsive family protein	2.35
	BF277659	ERF4 (ETHYLENE RESPONSIVE ELEMENT BINDING FACTOR4)	1.33
	TC74237	ERF3 (ETHYLENE RESPONSIVE ELEMENT BINDING FACTOR3)	1.43
	TC61850	EIN2 (ETHYLENE INSENSITIVE 2)	1.39
	TC63148	RTE1(REVERSION-TO-ETHYLENE SENSITIVITY1)	1.93
	TC77973	ethylene-responsive protein -related	1.35

Table 3.3, cont.

2 DPA	DN817447	ABI5 (ABA INSENSITIVE 5)	1.62
	TC69035	AtRAB1B (*A. thaliana* responsive to abscisic acid 1B)	1.45
	TC75959	BRH1 (BRASSINOSTEROID-RESPONSIVE RING-H2)	1.79
	TC65219	gibberellin-regulated family protein	1.82
	AY072821	RDL (RD22-like) (*Gossypium hirsutum*)	2.14
	TC73227	sucrose synthase	1.56
	BG445906	CESA8 (CELLULASE SYNTHASE 3)	1.32
	AW667548	AtMYB60	1.32
	TC59222	AtMYB73	1.37
	TC61396	AtMYB123, TT2 (TRANSPARENT TESTA 2)	1.39
	TC62952	AtMYB3	2.02
	TC75295	dormancy/auxin associated family protein	1.72
	TC65241	IAA29 (indoleacetic acid-induced protein 29)	1.60
	TC77973	ethylene-responsive protein -related	1.40
7 DPA	DN817447	ABI5 (ABA INSENSITIVE 5)	1.55
	TC65219	gibberellin-regulated family protein	1.48
	DT461247	peroxidase	1.44
	TC76863	peroxidase	1.65
	TC61410	AtGRF8 (GROWTH-REGULATING FACTOR 8)	1.63
	TC66179	AtADF (ACTIN DEPOLYMERIZING FACTOR 1)	1.43
	TC78808	ARPC3 (actin-related protein C3)	1.44
	TC75953	AtMYB60	2.05
	TC76829	auxin-responsive family protein	1.54
	CO098023	auxin-responsive protein	2.13
	TC72454	putative auxin growth promotor protein	1.55
	TC60797	ethylene-responsive protein	1.51
	TC61850	EIN2 (ETHYLENE INSENSITIVE 2)	1.56
	TC60226	CYP77B1 (cytochrome P450, family 77)	1.99
	TC67316	CYP81D8 (cytochrome P450, family 81)	1.51
	TC67461	CYP78A7 | CYP78A7 (cytochrome P450, family 78)	1.64
	TC64773	GL2 (GLABRA 2)	1.37
	TC65599	microtubule motor	1.65

FC: fold change

tubulin, and fiber protein E6. Interestingly, the genes encoding plant hormone related proteins such as auxin responsive factors, ethylene responsive factors, and gibberellin-regulated family proteins have been identified in the epidermal layers or fiber cells of cotton ovules. This result is consistent with a significant increase in the percentage of proteins involved in hormone signaling pathway in GH_TMO EST library (Yang et al., 2006). In addition, sucrose synthases, fiber proteins, and MYB transcription factors were repeatedly shown in the epidermis of cotton ovules during fiber initiation. Sucrose synthase is a well-known "fiber-related" protein which is localized in the fiber cells during fiber development and its suppression can cause repression of cotton fiber initiation and development (Nolte et al., 1995; Ruan and Chourey, 1998; Ruan et al., 2001; Ruan et al., 2003). Interestingly, dehydration induced protein RD22-like (*GhRDL*) was significantly up-regulated in the epidermal layers of 0 DPA ovules. RD22-like 1 in *Gossypium arboretum* (*GaRDL1*) has been studied (Li et al., 2002a; Wang et al., 2004). *GaRDL1* was highly enriched in developing fiber cells and the promoter of *RDL* in A-genome (*GaRDL1)* fused with a β-glucuronidase (GUS) showed trichome-specific expression in *Arabidopsis* leaves (Wang et al., 2004).

At 2 DPA, *TRANSPARENT TESTA 2* (*TT2*) was detected in epidermis. TT2 is a MYB transcription factor that is necessary for proanthocyanidin biosynthesis in *Arabidopsis thaliana* seed coat and associates with *TTG1* (Baudry et al., 2004). At 7 DPA, *GL2* (*GLABRA2*) is found in the fiber tissue at 7 DPA. *GL2* is a homeodomain protein that affects epidermal cell identity and down regulates seed oil content (Hülskamp, 2004;

Ishida et al., 2007). Their expression profiles were well confirmed by quantitative RT-PCR with *HISTONE H3* as a control (Figure3.3). Leaf and petal tissues were added to filter out the genes with high expression in non-fiber tissues.

Quantitative RT-PCR analysis confirmed the expression profiles of differentially expressed genes (Figure 3.3, Table 3.4). A ring zinc finger protein (AI730621) and fiber protein 37 (Fb37) were highly abundant in the epidermis of ovules at all four stages (-2 DPA, 0 DPA, 2 DPA, and 7 DPA) with their highest peak at -2 DPA (Fig 3.3A, B). The expression pattern of a transcription factor (TC74600) was mainly in the epidermis at -2 DPA (Figure 3.3C). After 0 DPA and 2 DPA, the expression of this transcription factor was higher in the inner ovules. On the other hand, the transcriptome level of *GhMyb112b* (TC75739) in the epidermis was increased to 2 DPA (Figure 3.3D). The gene encoding RD22-like protein (*GhRDL*) was highly expressed in the epidermis at 0 DPA and its expression level went up until 7 DPA (Figure 3.3E). Another differentially expressed gene, ABC transporter (TC75912) showed its highest expression in the epidermis at 0 DPA, then decreased afterwards (Figure 3.3F).

GO classification of differentially expressed genes

To further investigate gene regulation during fiber development, gene ontology classification of differentially expressed genes was performed in molecular function. At -2 DPA, the category of structural molecule activity was significantly overrepresented in inner ovule tissues compared to gene classes represented by the entire microarray total set

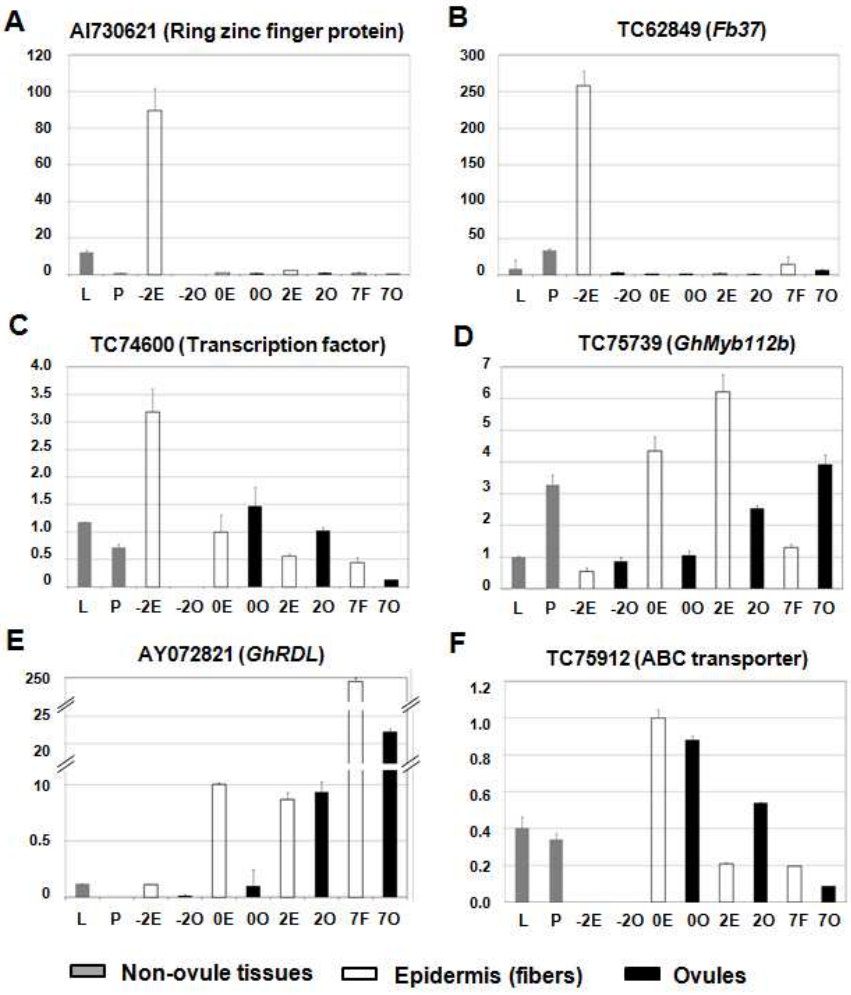

Figure 3.3. Expression patterns of differentially expressed genes.

Gene expression was analyzed in ten tissues. L: leaves; P: petals; -2E, 0E and 2E; epidermis at -2 DPA, 0 DPA and 2 DPA respectively. -2O, 0O , 2O and 7O; inner ovules at -2 DPA, 0 DPA, 2 DPA and 7 DPA respectively. 7F; fibers at 7 DPA. **(A)** ring zinc finger protein (AI730621). **(B)** fiber protein 37 (TC62849). **(C)** transcription factor (TC74600) **(D)** *Ghmyb112b* (TC75739), **(E)** *GhRDL* (AY072821). **(F)** ABC transporter (TC75912).

Table 3.4. Primers used for quantitative RT-PCR

EST ID	Quantitative RT-PCR Primers used
AI730621	F: CTT TTC TGC TGG CCT TGC TT
	R:GCA GAC AGT GTG AAA TTC CCG
TC62849	F:CCC ACC TCT TTG AAA CTC TTG GCT
	R: CAA TTG TTC GCA CAC AGC TTC GGT
TC74600	F:GTG ATG ATG TTA CCA GCC CCA
	R:AGC ACC AGT TGC GGA TAC ACA
TC75739	F: AAC CGC AAT GCC TCG ATG TAC TCA
	R:ACG TCG ATG TTG GAG ATT GGA GGT
AY072821	F:ATC AGC ATG GAA CCC TAA GCA
	R:GGA CCC AAA CAA TGT GAT CCC
TC75912	F:CCA TGG CTG CAA TGG ATT AAG
	R:CCA AGT GCC CAA AAA CAT CTG
TC58926	F: GGC AGC AAC ATC TTT GCT GAT GGA
	R: AAG CAA TGC TGC TTT CAC CAC CTG
TC60696	F: TAA GCC TGC GAG CTT CGG TTC TAA
	R: GGA ATA ATT TCA GGT TCG GCG CGT
TC69088	F: GCA CTA GCA ATG CCA AGG GTC AAT
	R: TTG TCA CAG CAG AGG AGA AGC TCA
CO094001	F: GAC ACC ATG CAG GTC ATT CAT CCA
	R: GCG ACA GCT GCT GCT GAT AAA GTT
DT543487	F: TTC AGC ACA CTC CTT ATG CCT CTC
	R: TCT CTT CAA AGA GGA AGG TGC CCA
TC73047	F: ATG GAG GCA ATG GAG GTT GCA ATG
	R: TGG AAG AAG AAA TGG CTG CGT TGG
TC68074	F: TCT TTC TGC GTT CCA AGC CTT TGC
	R: ATC TCC TTC CGG TGG CAT CAT TCA
TC73631	F: AGC CTT AAG TTG GGT CGA CGT GAT
	R: ATT CCA TGG CCA AGA CTT GCA ACC
TC76555	F: ATG CGA AGA TTT GGA GAG CAA CGG
	R: ATC CCT TCC GCC ACT TAT AAG GCT
TC60063	F: TTT GCG ACT TCT TGA GAG GAG GGA
	R: TGG GTA CCT GGC CTT TAC ACT GTT
TC73445	F: ACA ACC CTG ATA ACG ACA CCA CCA
	R: CCA ATC ATG GCA CGA CAA CCA CTT
TC63972	F: ACT TGT CGG TGG TTC AGA GCT TGT
	R: TAT TTG CAG GTG ATG GAC GAG GCT
TC67191	F: TAT CGC ATG ACG GAC CTG GTT TCT
	R: AGG GAG ACA GAT CAC TTT GGC GTT
TC76014	F: GCA TCC ACC TTG CAA CAA CGG TAA
	R: TCG TGC TTA CAC CTG AGG AGC TTT
TC63385	F: AGA GGG CCC GCA TTG TAT CAA CTA
	R: CCT GTT GCC TGT TGT TGC TGA TGT
TC69040	F: AGG CTC GAT GGA AGA GTC GTT GAT
	R: TGC GAC TTC TTG ACC TTG GAA CTG
TC69095	F: TAT CTG GGT CCT GTC GTC GTC TTT
	R: ATT GGA ATT CCC ACT GGA CGA TGC

F: forward primer, R: reverse primer

gene ontology in molecular function (Figure 3.4B). Before or right after the cotton ovule receives a signal to promote fiber initiation, ovules may prepare for massive transcriptional regulation and cell wall expansion, making lots of structural molecules including structural constituents of ribosomes and tubulins in the inner ovule tissues. When ovules are fertilized and start to make fiber initials on their epidermal layers, genes for structural molecule activity seem to move from ovules to fiber initials and the overrepresentation goes away (data not shown). This may be related to dynamic changes in size and structure in fiber initials (Ramsey and Berlin, 1976; Ryser, 1985; Berlin, 1986; Ryser, 1999). *In vitro* ovule cultures, it has been shown that the nucleoli are active in the formation of pre-ribosomal particles for ribosome formation and protein synthesis and a large proportion of ribosomes is probably produced for the rapid elongation (Meinert and Delmer, 1977; Berlin, 1986). In the genes which were up regulated in the epidermis at -2 DPA, none of the molecular function categories statistically differ with p value < 0.05 (Figure 3.4A). At 0 PDA and 2 DPA, there was no over- or under-representation of functional categories in the inner ovules as well as fiber initials.

At 7 DPA, DNA or RNA binding proteins were overrepresented in the inner ovules whereas genes for transporter activity were highly abundant in the fiber tissues (Figure 3.4A, B).

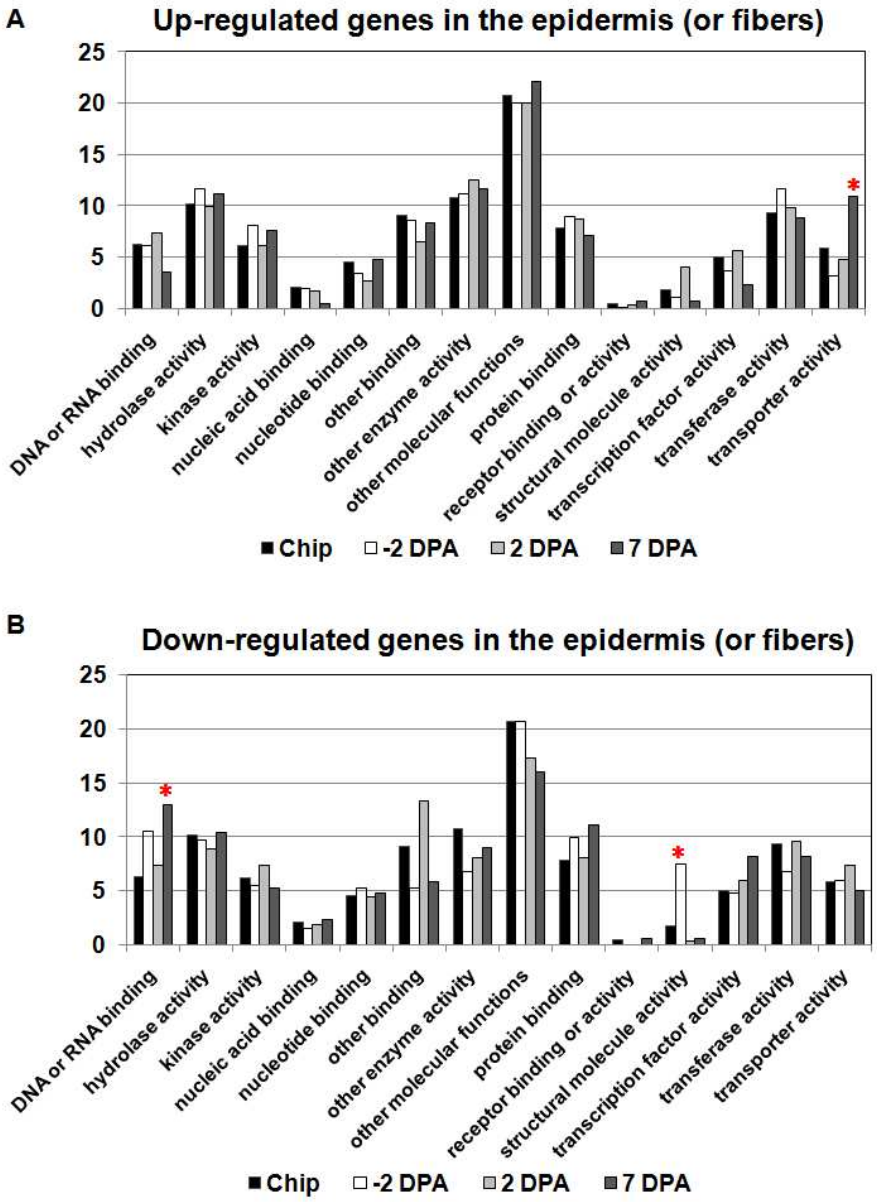

Figure 3.4. Gene ontology analyses of differentially expressed genes.

Molecular functions of up-regulated genes **(A)** and down-regulated genes **(B)** in the epidermis (or fibers). *significantly overrepresented category with p value > 0.05.

Clustering analysis of differentially expressed genes

The differentially expressed genes in 4 developmental stages including -2 DPA, 0 DPA, 2 DPA, and 7 DPA were clustered by K-means clustering with Euclidean. This makes 6 clusters based on their expression profiles (Figure 3.5, Table 3.5). The summary of their expression profiles is shown as a heat map in Figure 3.5A.

Genes in cluster 1 showed only subtle changes during fiber development whereas genes in other clusters had dramatic expression differences in each developmental stage (Figure 3.5B). Cluster 1 includes the genes which have the transferase activity such as S-adenosylmethionine adenosyltransferase (SAM). SAM catalyzes the formation of S-adenosylmethionine from methionine and ATP and is known to be a major methyl group donor in living cells (Lamblin et al., 2001). This enzyme is involved in the biosynthesis of ethylene and essential for cell metabolism including seed germination and seedling establishment (Gallardo et al., 2002). Thus, it is possible that there is an active cell metabolism throughout fiber development.

Gene expression was high in the epidermis at -2 DPA, went up at 0 DPA, and then decreased after that in cluster 2. This class mainly includes the genes involved in kinase activity as well as the genes encoding cotton boll abscission zone (AI055589), shaggy-like kinase encoded by *BIN2* (AI725469), and auxin response factor 36 (TC68975) fall into this cluster. Cluster 3 showed the highest expression level at -2 DPA and went down at 0 DPA. Genes with hydrolase activity were found in this class. Cluster

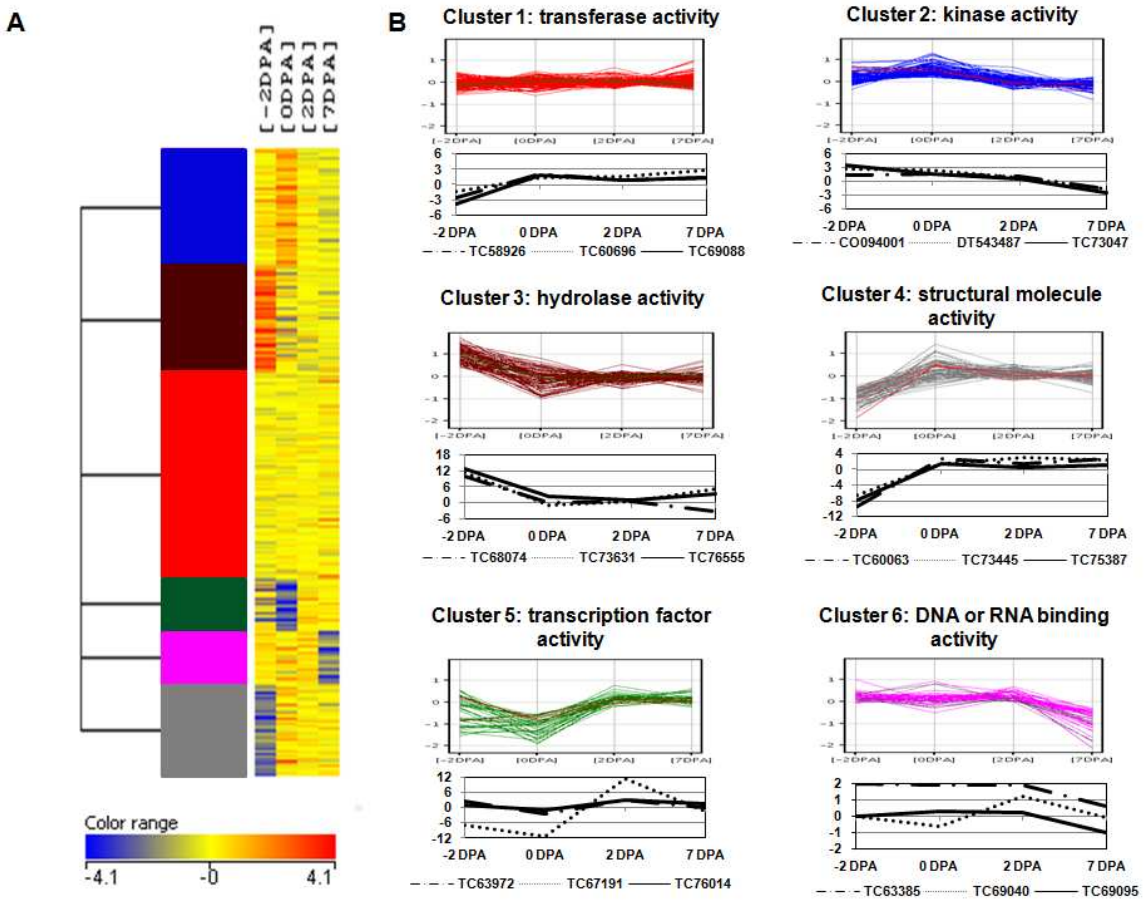

Figure 3.5. Clustering analysis of differentially expressed genes at -2 DPA, 0 DPA, 2 DPA and 7 DPA.

(A) A heat map of K-means clustering.
(B) Expression profiles of six clusters with qRT-PCR verification.

Table 3.5. List of genes found in each cluster

Cluster	EST ID	Tentative annotation
1	AI727497	SAM1 (S-adenosylmethionine synthetase 1)
	BE052240	SAM-2 (S-adenosylmethionine synthetase 2)
	TC58926	SAM-2 (S-adenosylmethionine synthetase 2)
	TC60696	pyruvate kinase, putative
	TC69088	ATMPK2 (mitogen-activated protein kinase homolog2)
	TC70118	ATHST (prenyltransferase)
	TC71501	protein kinase family protei
	TC72046	adenylylsulfate kinase, putative
2	AI055589	serine/threonine protein kinase, putative
	AI725469	BIN2 (BRASSINOSTEROID-INSENSITIVE 2)
	CO094001	ribitol kinase, putative
	DT543487	ROPGEF1 (KINASE PARTNER PROTEIN-LIKE)
	TC63543	ribitol kinase, putative
	TC73047	SRPK4 (SER/ARG-RICH PROTEIN KINASE 4)
	TC76694	phosphatidylinositol-4-phosphate 5-kinase family protein
	TC77576	leucine-rich repeat transmembrane protein kinase, putative
3	BM359206	ZFN3 (ZINC FINGER PROTEIN 3)
	TC59978	zinc finger (AN1-like) family protein
	TC78190	calcineurin-like phosphoesterase family protein
	TC77847	ATACP5 (acid phosphatase 5)
	TC70366	HDA2 (histone deacetylase 2)
	TC73631	BGAL13 (beta-galactosidase 13)
	TC76555	FAC1 (EMBRYONIC FACTOR1)
4	TC60063	ATMAMI (membrane-associated mannitol-induced)
	TC60196	60S ribosomal protein L10A (RPL10aC)
	TC73239	TUA4 (tubulin alpha-4 chain)
	TC73445	60S ribosomal protein L8 (RPL8C)
	TC75387	PS6 (RIBOSOMAL PROTEIN S6)
	TC64286	r ibosomal protein L12 family protein

Table 3.5, cont.

	TC68777	KOW domain-containing protein
5	TC63972	zinc finger (C2H2 type) family protein
	TC67191	AT-HSFB2A (heat shock transcription factor B2A)
	TC74741	ABF2 (Abscisic acid responsive elements-binding factor2)
	TC75739	AtMIXTA/AtMYB16 (MYB domain protein 16)
	TC76014	transcription factor
6	TC63385	ATSWI3C (Arabidopsis thaliana switching protein 3C)
	TC66508	auxin/aluminum-responsive protein, putative
	TC66510	auxin/aluminum-responsive protein, putative
	TC66798	ATP binding / damaged DNA binding
	TC67889	HAT9 (homeobox-leucine zipper protein 9)
	TC69040	ATSC35 (arginine/serine-rich splicing factor 35)
	TC69095	bZIP transcription factor family protein
	DT458844	IAA29 (indoleacetic acid-induced protein 29)
	TC65219	gibberellin-regulated family protein
	TC79957	gibberellin-regulated family protein

3 also includes several genes encoding alcohol dehydrogense 2 (*ADH2*; TC66247, TC66349, and TC66366), and fiber protein 37 (*Fb37*; TC62849). Genes encoding alcohol dehydrogenase have shown the fiber preferential expression pattern in suppression substractive hybridization (SSH) (Gao et al., 2007). *ADH* may play a role during fiber initiation, functioning in various metabolic processes in the cells.

Cluster 4 had a highest expression in the epidermis at 0 DPA. Several genes that have the structural molecule activity were identified in this cluster. The structural molecules such as ribosomal proteins and tubulin factors seem to work in the inner ovules at -2 DPA and move to the fiber initials afterwards. This data is consistent with the previous finding of preferential accumulation of tubulin in the fiber cells during cotton fiber cell elongation (Dixon et al., 1994; Ji et al., 2002; Li et al., 2002b; Shi et al., 2006) . Cluster 4 also contains the gene encoding BES/BZR1 homolog protein 3 (TC80304).

The expression profile of cluster 5 showed the lowest expression in the epidermis at 0 DPA, meaning that the genes in this class are abundant in the inner ovules at this stage. The expression is increased after 0 DPA. This includes the genes encoding several transcription factors (TC67191, TC75739, TC76014) and an abscisic acid responsive elements-binding factor 2 (TC74741) as well as a gene encoding a germin-like protein (TC72843). The germin-like protein has been reported to be accumulated during fiber elongation (Kim et al., 2004c) which matches with this expression data.

The last class, cluster 6 had up-regulation in the epidermis during fiber initiation stage and then expression showed in the ovules at 7 DPA. This class contains genes involved in

the function of DNA or RNA binding activity. It is possible that transcriptional regulation becomes active during fiber elongation requiring many key players with DNA or RNA binding activity. In addition, two gibberellin regulated genes (TC65219 and TC79957) and a sucrose synthase (*Sus*) (TC73327) have been found in this cluster. *Sus* is the one of the most well-known fiber-associated genes in cotton (Nolte et al., 1995; Ruan and Chourey, 1998; Ruan et al., 2001; Ruan et al., 2003).

Functional analysis of *GhRDL* in *Arabidopsis* trichome development

Dehydration induced protein RD22-like in *Gossypium hirtutum* (*GhRDL*) has shown the fiber preferential gene expression during fiber initiation and elongation and selected for functional analysis (Figure 3.3). *Arabidopsis* can be a useful model system for understanding fiber cell initiation because cotton fibers, as seed trichomes, have many similarities with leaf trichomes in *Arabidopsis*. The promoter of *GaRDL1* was already showed trichome-specific expression in *Arabidopsis* (Wang et al., 2004).

GaRDL1 shares 97% identity with *GhRDL* in amino acid whereas it shows some deletions compared to *AtRD22* (Figure 3.6). *RD22* in *Arabidopsis* is induced by ABA, but its seed-specific expression is not regulated by ABA(Yamaguchi-Shinozaki, 1993). To see if *GhRDL* is functioning in trichome development in *Arabidopsis*, the overexpression study was performed with full length cDNA of *GhRDL* (Figure 3.7A). After transformation into *Arabidopsis* and selection for transformants, the transgenic plants harboring *35S:GhRDL* were genotyped with a pair of gene specific primers used for

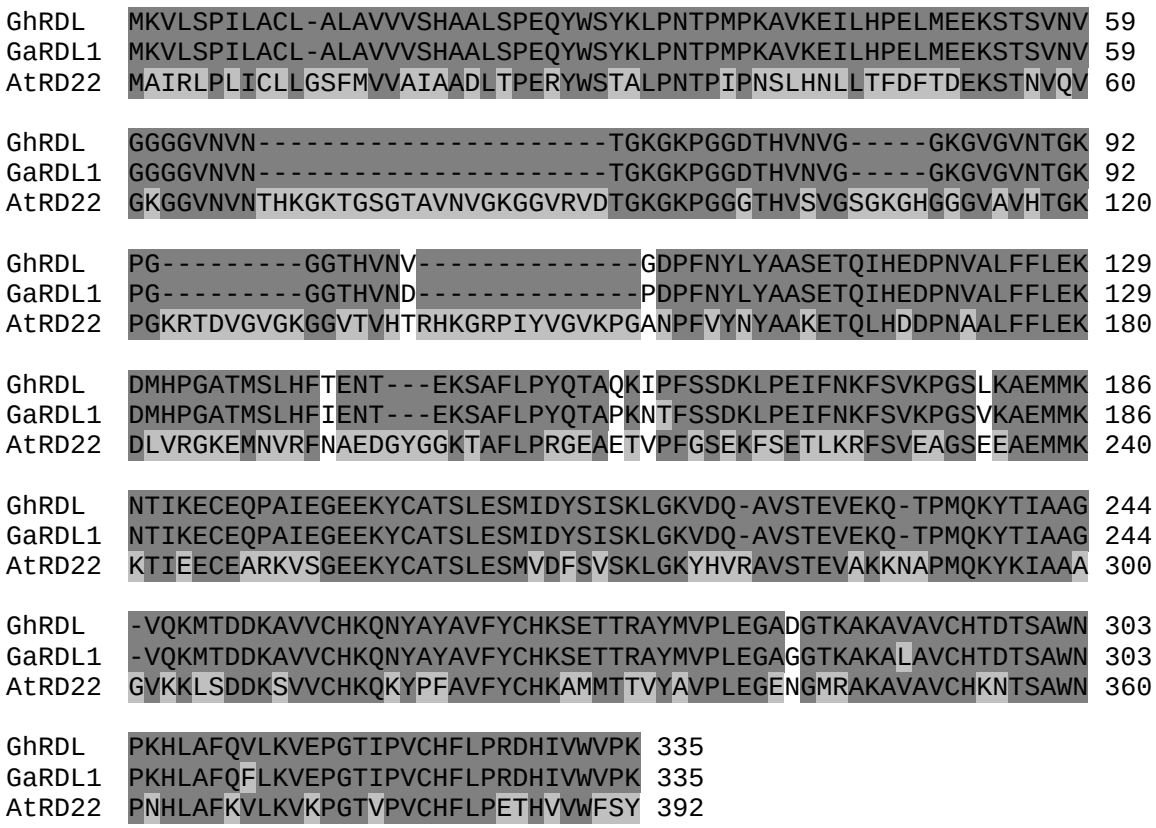

Figure 3.6. Sequence alignment of RDL and RDL-related genes.

Accession numbers: GhRDL (AAL67991, *Gossypium hirsutum*), GaRDL (AAT66912, *Gossypium arboreum*), AtRD22 (Q08298, *Arabidopsis thaliana*)

cloning. The insert was amplified from five transgenic lines in the same size with that in the binary vector (Figure 3.7B), confirming that they contain the *GhRDL*. The transcriptome of *GhRDL* was assayed with quantitative RT-PCR with the gene specific primers used for microarray data validation to see if the transgene is actively transcribed

in the transgenic lines. Transgenic line 4 showed highest expression among 5 individual lines (Figure 3.7C).

The phenotypes of transgenic plants were observed with light microscope. The plants transformed with *35S:GhRDL* showed a significant increase in the number of trichomes on their leaves (Figure 3.8A and C) compared to the wild type *Arabidopsis* (Figure 3.8B and D) (Table 3.6). However, the trichomes on stems and floral organs of transgenic plants (Figure 3.8E and G) had no significant changes compared to that of the wild type *Arabidopsis* (Figure 3.8F and H). The *GhRDL* did not complement glabrous phenotype in *gl1* (CS1644) mutant, either (data not shown). The complementation of *gl1* may require introns of *GhRDL* as it did with *GaMYB2* (Wang et al., 2004). It is also possible that *GhRDL* plays a role in the downstream of *GL1* during trichome development. This transgenic plant is derived from the 35S:GhRDL-4 plant which showed the highest expression of *GhRDL* among 5 transgenic lines (Figure 3.7C). It is known that the L1 box and MYB motif in the promoter of *GaRDL* contribute to the trichome expression in *Arabidopsis* (Wang et al., 2004). Thus, it is very likely that RDL has a significant role in the trichome development pathway, implying a possible function in cotton fiber development pathway.

To gain more data for phenotypic changes caused by *GhRDL*, I examined leaves and seeds of transgenic *35S:GhRDL* plants using scanning electron microscope (SEM).

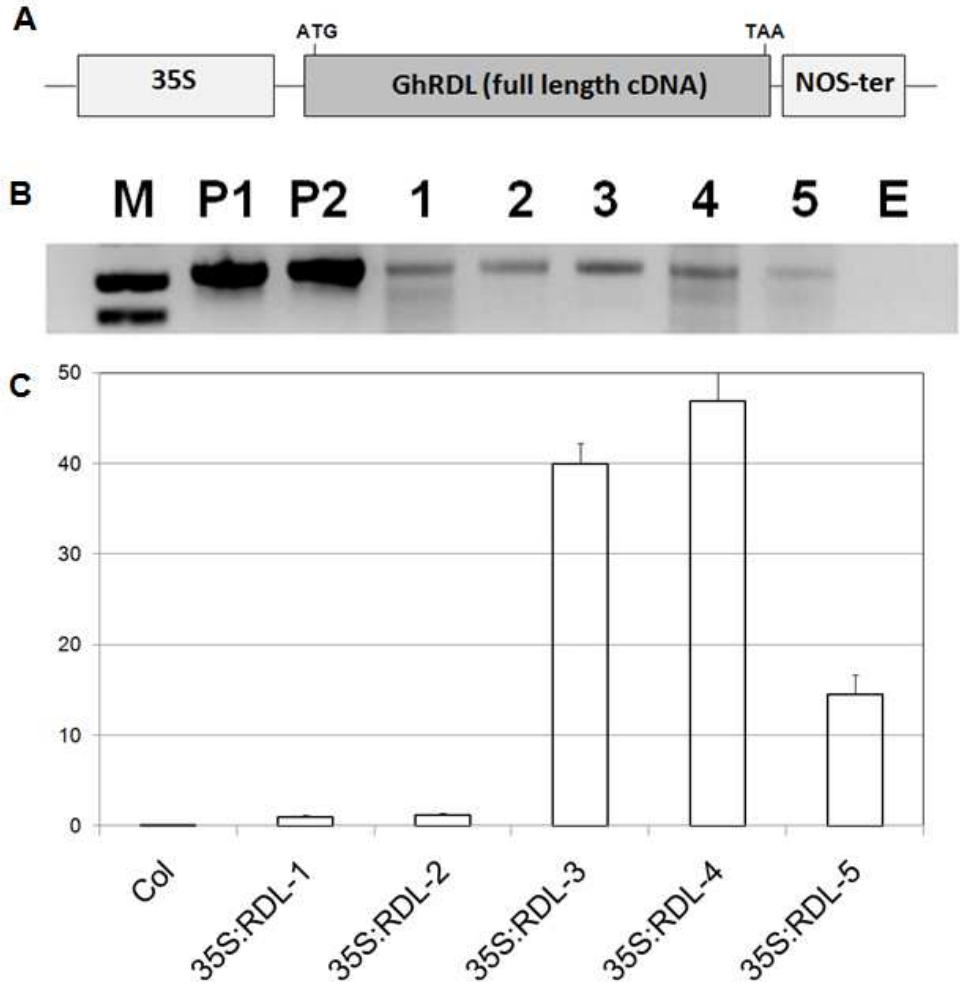

Figure 3.7. Overexpression of *GhRDL* in *Arabidopsis*

(A) Overexpression vector harboring *GhRDL*
(B) Genotyping of transgenic *Arabidopsis.* **M**; 1Kb marker. **P1, P2**; inserts from the binary vector. **1-5**; transgenic lines haboring *35S:GhRDL* . **E**; Transgenic line transformed with empty vector.
(C) Gene expression of *35S:GhRDL* in transgenic plants

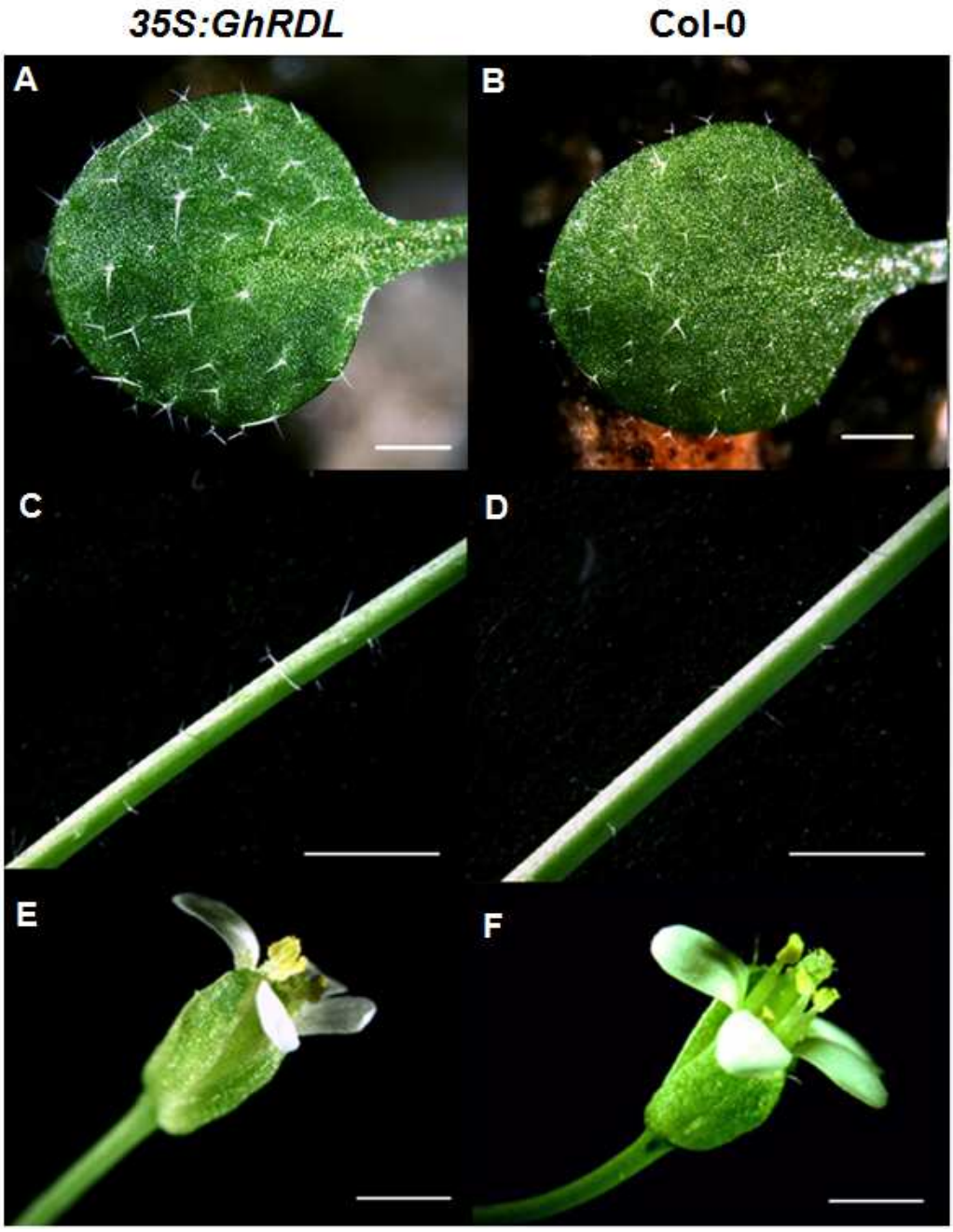

Figure 3.8. Trichome phenotypes of transgenic plants bearing *35S:GhRDL* **(A), (C), and (E)** and wild type *Arabidopsis*. **(B), (D), and (F).** The scale bar in each figure is 2 mm.

Table 3.6. Average number of trichomes on fully expanded first leaves

Genotypes	Number of trichomes/leaf[a]	Number of leaves examined
Wild type Columbia	30.9 ± 0.8	20
35S:RDL	38 ± 1.2	11

[a]Mean ± s.e.

The structure of leaf trichomes in the transgenic plants was similar to that in wild type *Arabidopsis* (Figure 3.9). However, an interesting phenotype was observed with the seeds of transgenic plants. *35S:GhRDL* in the wild type genetic background promoted unbranched seed hairs and the hairs in the transgenic plants were phenotypically different from trichomes on the (Figure 3.10). This showed an unbranched hair structure like cotton seed fibers on SEM analysis and its length was up to ~ 0.2 mm long (Figure 3.10). Among 112 seeds examined by SEM, 29 seeds had at least one trichome (about 26%) and some had more than 20 seed hairs.

Despite similarities between fiber development in cotton and trichome development in *Arabidopsis*, these two cell types differ in terms of endoreduplication and in branch formation (Hülskamp et al., 1994; Marks, 1997; Hülskamp and Schnittger, 1998; Hülskamp, 2004; Lee et al., 2007). In general, *Arabidopsis* does not have hairs or trichomes on its seeds. Therefore, seed hairs induced by *GhRDL* imply that this cotton gene mimics cotton fiber development or trichome development in the seed epidermis of

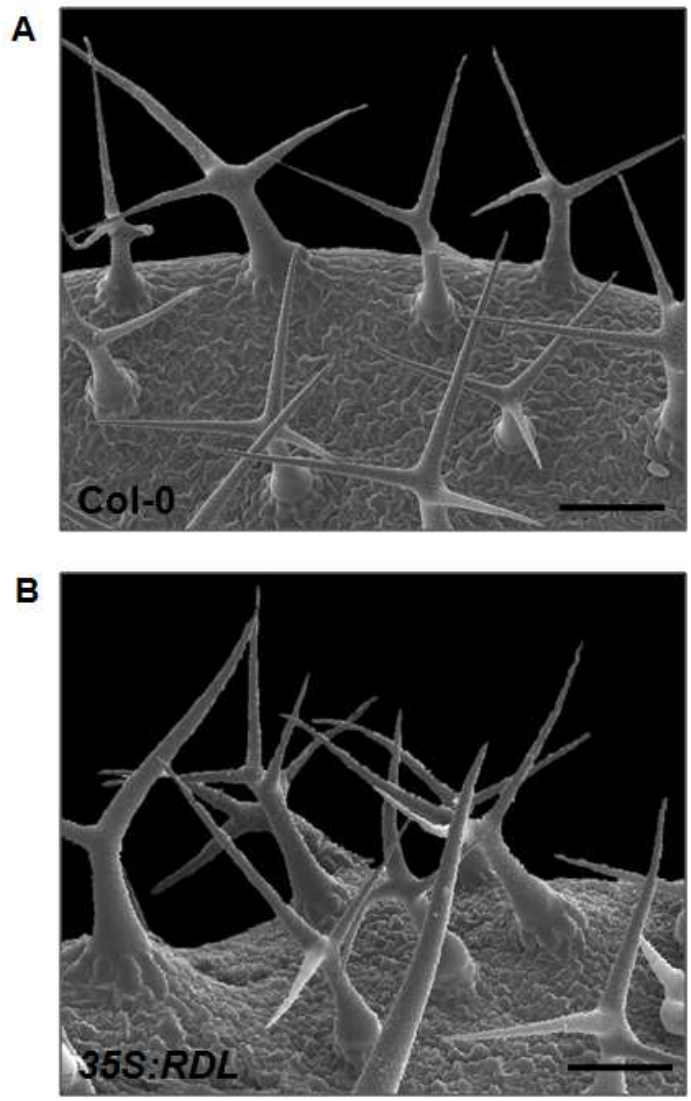

Figure 3.9. Scanning electron micrographs of leaf trichomes of the wild type *Arabidopsis* **(A)** and the transgenic *Arabidopsis* bearing *35S*:*GhRDL* **(B)**. The scale bar in each figure is 100 μm.

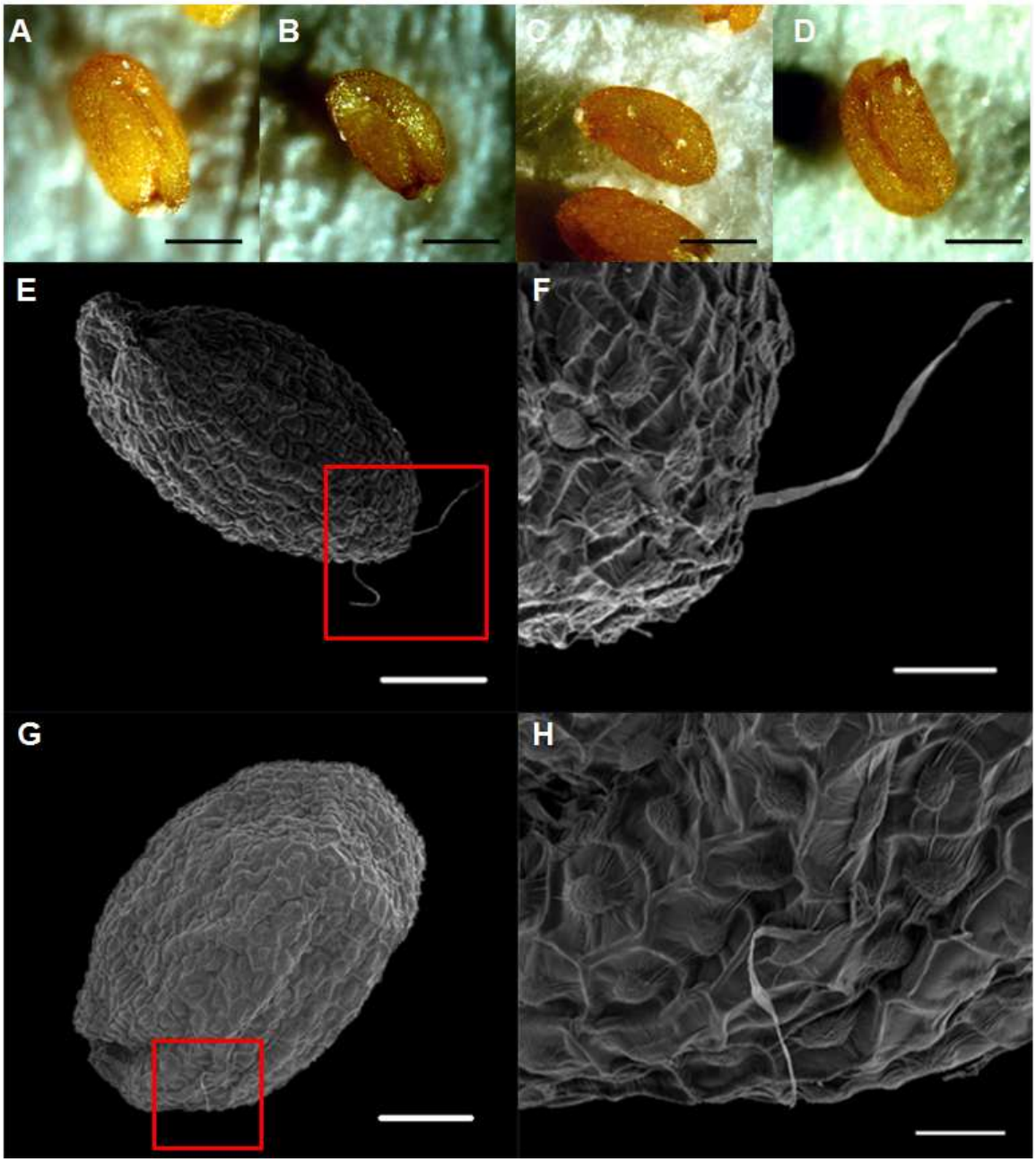

Figure 3.10

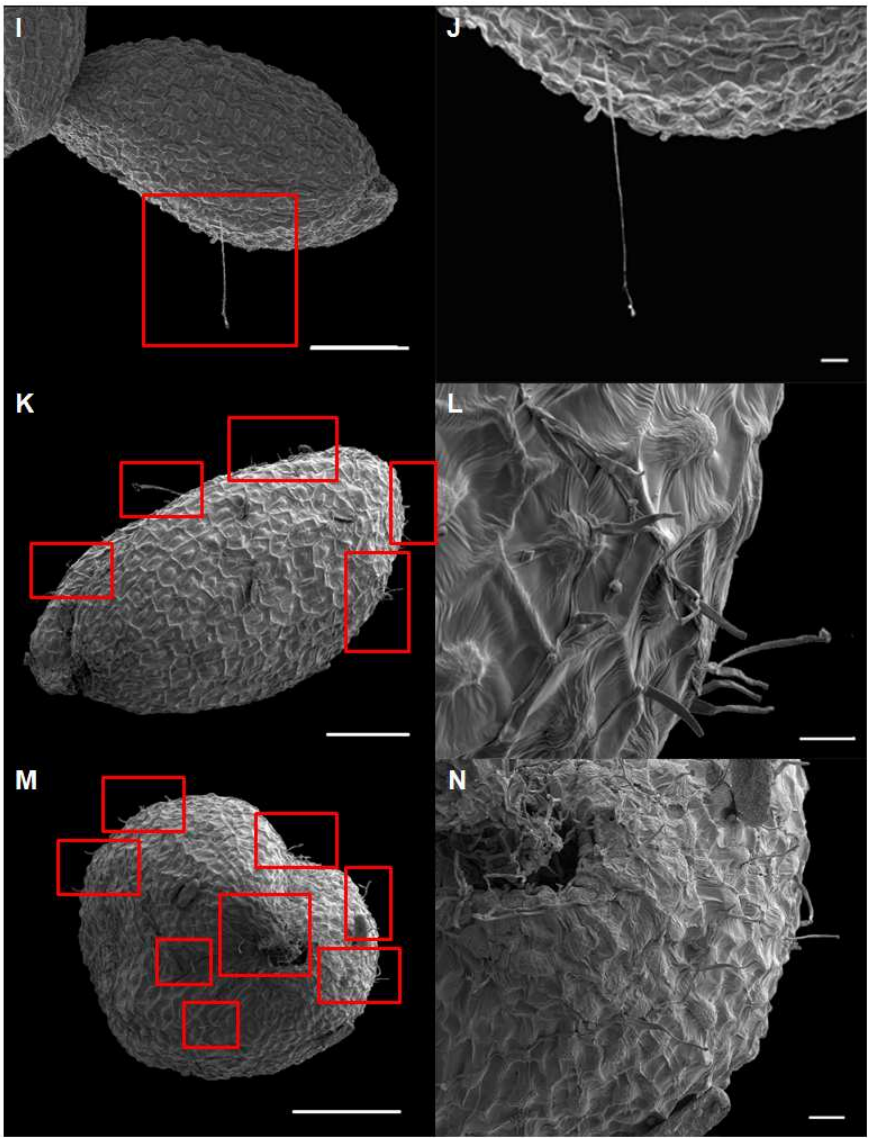

Figure 3.10. Dissection microscope pictures ((**A**) to (**D**)) and scanning electron microscopic pictures ((**E**) to (**N**))of depicting the seed hairs of transgenic *Arabidopsis* bearing *35S*:*GhRDL*.

Enlarged view of seed hairs in **(F)**, **(H)**, **(J), (L), and (N)** are from **(E)**, **(G)**, **(I), (K), and (M)**. The scale bar in (A), (B), (C), and (D) is 1mm, in (E), (G), (I), (K), and (M) is 100 µm and in (F), (H), and (J) is 20 µm, and (L) and (N) is 10 µm.

Arabidopsis thaliana. It also suggests that *GhRDL* may have a significant function during cell differentiation and fiber development in cotton.

Localization of *GhRDL* in Arabidopsis

The promoter of *GhRDL* was obtained by genome walking and fused with β-glucuronidase (GUS) reporter gene to study the regulatory mechanism of *GhRDL*. The fragment includes from ~450 bp upstream of the transcription initiation site to the second intron (Figure 3.11A). *GhRDL* contains several regulatory elements in its promoter region (Figure 3.11B). Interestingly, CE1 regulatory element is involved in the regulation of a ABA responsive gene *HVA22* and DAE motif is known as GA-responsive promoter, suggesting hormonal regulation of gene. GT motif at -312 is the site for the transcription factor of *Adh*. The transformants bearing this construct showed the GUS activity in their seeds (Figure 3. 11C). However, no GUS activity was detected in leaves, stems and roots (Figure 3.11C). This data suggests that *GhRDL* is involved in seed development in *Arabidopsis* and can be also associated to fiber development in cotton.

DISCUSSION

Investigation of "fiber-associated" genes that are mainly involved in fiber initiation was challenging due to technical difficulties of getting only fiber initials from whole ovule tissues. Laser capture microdissection (LCM) is now widely used to solve the limitation

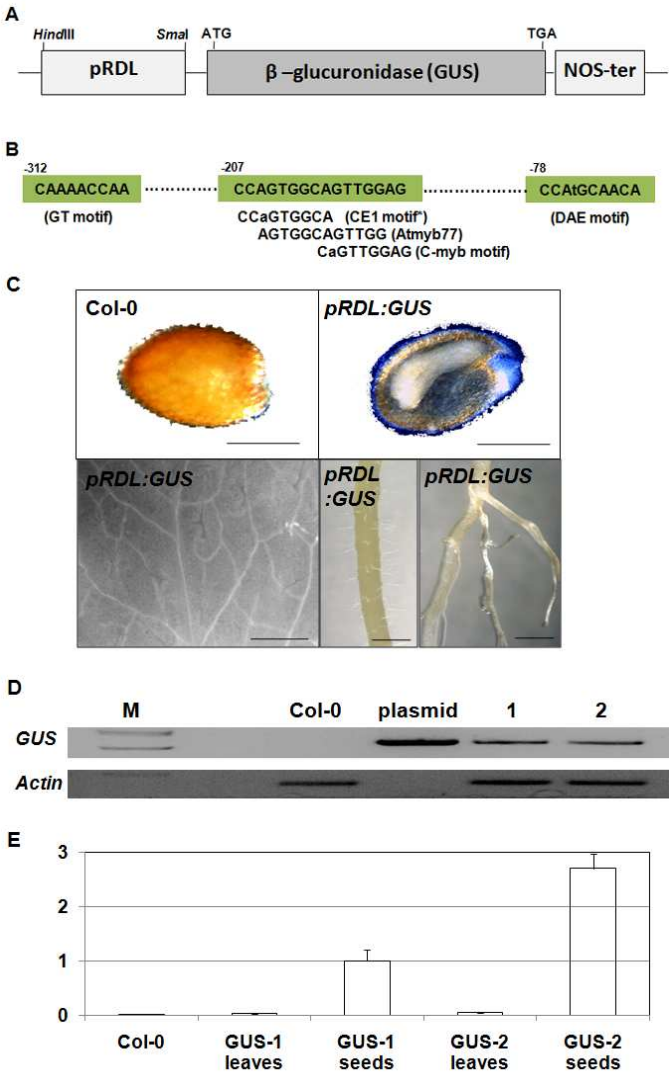

Figure. 3.11. Localization of *GhRDL* in *Arabidopsis*

(A) The construct for GUS assay.
(B) Regulatory element in the promoter region of *GhRDL.*
(C) GUS staining of the transgenic seed expressing *pRDL:GUS*. The scale bar is 0.5 mm in the top panel and 0.2mm in the bottom panel.
(D) Genotyping of transgenic *Arabidopsis.* **M**; 1Kb marker. **Col-0**; Arabidopsis thaliana, Columbia-0. **plasmid**; binary vector harboring *pRDL:GUS* . **1-2**; transgenic lines harboring *pRDL:GUS*.
(E) Gene expression of *pRDL:GUS* in transgenic plants.

on tissue collection (Jin et al., 1999; Matsunaga et al., 1999; Ohyama et al., 2000; Nakazono et al., 2003; Rook et al., 2004; Casson et al., 2005; Angeles et al., 2006; Cai and Lashbrook, 2006). In this study, I used LCM to collect the epidermis from which fiber initials are derived. Global gene expression studies with the laser microdissected epidermal cells and inner ovule tissues identified several sets of "fiber-associated" genes. In each developmental stage, different set of genes were regulated in the epidermis and inner ovules. In the inner ovules at -2 DPA, genes involved in structural molecule activity seem to prepare for massive metabolic pathways required for fiber initiation (Figure 3.4B). During fiber elongation, genes involved in DNA or RNA binding activity were highly abundant in the inner ovule, suggesting vigorous transcriptional regulation (Figure 3.4B). Genes with the transporter activity were enriched in the fiber cells at 7 DPA (Figure 3.4A). This may suggest that communication is an important factor in the fiber cells to accommodate necessary materials for fiber elongation. Clustering analysis also indicated that there is temporal and spatial gene regulation during fiber development and its expression profile differs in various gene categories.

Application of LCM in cotton fiber has been reported in 2006 (Wu et al., 2006) and applied for cDNA microarrays to profile gene expression differences between the fiber initial cells and non-fiber epidermal pavement cells (Wu et al., 2007b). The most highly up-regulated genes in cotton fiber initials include cytochrome c oxidase, putative reverse transcriptase, ubiquitin-protein ligase, and NADH dehydrogenase (Wu et al., 2007b). Although the differentially expressed genes cannot be compared with those I

found side by side because of different cell types used (fiber initial vs. non-fiber epidermis and epidermal layer vs. inner ovules), ~64% of the top 50 highly expressed genes in cotton fiber initial cells in their study were also up-regulated in the epidermis of cotton ovules during fiber initiation in this study.

Transcription profiling and ovule culture experiments of cotton have implicated plant hormones mediate cotton fiber development. With exogeneous auxin and gibberellin, unfertilized ovules were able to produce fibers in vitro (Beasley, 1971), suggesting auxin and gibberellins are crucial players to promote fiber cell development (Kosmidou-Dimitropoulou, 1986; Gialvalis and Seagull, 2001; Lee et al., 2007). Furthermore, *in vivo* assay revealed a spike of auxin levels in flower buds during fiber initiation (Guinn and Brummett, 1988). The enrichment of auxin transcripts in the epidermis was also detected in this study (Table 3.3). ATP binding cassette (ABC) transporters are involved in transporting diverse substrates including auxin across membranes. Asymmetrical localization of uptake and efflux carriers controls auxin gradients through auxin transporters (Luschnig, 2002). ABC transporters have been reported in the pilot microarray experiment (Lee et al., 2006) and also detected in the epidermis during fiber development (Figure 3.3).

In addition, apyrase (nucleoside triphosphate-diphosphohydrolase) genes are influenced by auxin and their expression is closely related with plant growth. Expression of two *Arabidopsis* apyrase (*APY1* and *APY2*) genes showed the highest expression in the tissues that accumulate high auxin levels (Wu et al., 2007a). Interestingly, a cotton

homolog of *APY* was low-abundant during fiber initiation but highly induced during fiber elongation (at 7 DPA) (data not shown), indicating that apyrase also plays an important role in fiber development.

It is expected to detect enrichment of trichome-related genes in the epidermis of cotton because their orthologous genes found in cotton. However, just a few genes showed differential expression in the epidermis of cotton. It can be because gene expression differences of these genes were not significant enough to detect in microarray level. It is also possible that some genes are not tissue-specific like *TTG1* in *Arabidopsis* (Zhang et al., 2003; Baudry et al., 2004; Humphries et al., 2005).

One of the "fiber-associated" genes was *dehydration induced protein RD22-like* (*GhRDL*). The expression of *GhRDL* was detected preferentially in the epidermis (Figure 3.3E). *RD22-like 1* in *Gossypium arboretum* (*GaRDL1*) has been studied in *Arabidopsis* (Wang et al., 2004). It was predicted that *GhRDL* has a significant role in the trichome development in *Arabidopsis* and in the fiber initiation in cotton because the promoter of *GaRDL* has already shown the trichome specific regulation (Wang et al., 2004) .

To elucidate the possible role of *GhRDL* in cotton fiber initiation, the construct of *35S:GhRDL* was introduced into glabrous *gl1* mutant and the wild type *Arabidopsis*. Although the glabrous phenotype in *gl1*mutant was not complemented with *35S:GhRDL*, overexpression of *35S:GhRDL* in the wild type promoted seed hair development as well as dense leaf trichome development, suggesting that *GhRDL* functions in the seed hair and trichome development pathway. Consistent with this, *pRDL:GUS* was localized in

Arabidopsis seeds. The expression of *RDL* may be mediated by plant hormones, ABA and GA through ABA-responsive CE1 motif and GA-responsive GT motif (Figure 3.11C).

It further implies the positive regulation of *GhRDL* in cotton fiber initiation because *Arabidopsis* trichome development is similar with cotton fiber development in many ways. Studying the relationship between *GhRDL* and previously studied fiber specific gene *GaMyb2* will be a noteworthy work to reveal the mechanisms of cotton fiber initiation. It is also possible that *GhRDL* is working with several MYB transcription factors as well as plant hormones which may be required for fiber development. Thus, discovery and investigation of interacting genes in the network of fiber initiation is necessary to understand the fundamental phenomenon of fiber development.

METHOD

Plant materials

Gossypium hirsutum L. cv. TM-1 (>S6 generation) was grown in a greenhouse at the University of Texas. Flower buds prior to anthesis (−2 DPA) were collected when the ovules were enclosed by squares of 1/4 ~ 1/3 inches in diameter. Flowers were tagged on the day of anthesis, and ovules were harvested at 0, and 2 DPA. Cotton bolls (TM-1) at 7 DPA were harvested for dissecting fibers. For each genotype, we used two biological pools, each with ten plants grown at similar stages. Leaves were harvested by pooling one

leaf from each of ten individual plants. Ovules or fibers were either dissected from five bolls collected in each of ten plants. The fresh tissues were frozen in liquid nitrogen and stored in a −70°C freezer, subjected to RNA extraction, or subjected to tissue fixation for laser capture microdissection. *Arabidopsis (Arabidopsis thaliana* (Columbia-0)) seeds were sterilized with 10% (v/v) house bleach for 10 min, followed by two washes in sterile water. Seedlings were germinated on plant germination medium (GM) in the growth chamber under continuous illumination at 22°C. Plants were transferred to soil (MetroMix 200, Sungro, Bellevue, WA) and grown in 16 h photoperiod at 22°C.

EST selection and 70 oligonucleotide design for microarray

A large set microarray was designed from several cotton EST libraries. This includes three sets of oligonucleotide probes (1,536, 12,006, and 9,629, respectively), making a total 23,171 of oligosnucleotides on a single chip .

The first set of 1,536 oligonucleotide probes was designed from a pilot set of cotton microarray in Chapter III (Lee et al., 2006). The second set of 12,006 oligonucleotide was designed from an exemplar sequence set which was chosen from a clustere set of unigenes by single-linkage clustering with BLASTN using Picky v1.0 (Chou et al., 2004; Udall et al., 2006; Udall et al., 2007). The 12,006 oligonucleotide probes include genes encoding a large number of transcription factors, and several thousand genes that had homology to *Arabidopsis* genes (Udall et al., 2007). The third set of 9,629 oligonucleotides probes was designed from CGI8 (http://compbio.dfci.

harvard.edu/tgi) that contained 55,673 unique sequences using Picky v2.0. This GH_TMO ESTs showed the accumulation of genes encoding putative transcription factors such as MYB and WRKY and genes encoding predicted proteins involved in auxin, brassinosteroid (BR), gibberellic acid (GA), abscisic acid (ABA) and ethylene signaling pathways (Yang et al., 2006). The probes that target the same genes as in the first two probe sets were excluded in the third set. This large set of cotton oligonucleotide microarray may include about 46–60% of the total genic diversity and most of the probes had a single target (Udall et al., 2007).

Microarray experimental design

For gene expression studies using a large set cotton oligo-microarray, 4 developmental stages were chosen. To study differential expression during fiber initiation, ovules at -2 DPA, 0 DPA, and 2 DPA were used. One of the fiber elongation stage tissues (7 DPA) was included. In each developmental stage, epidermis was separated from inner ovules and subjected to the hybridization. In addition, epidermis and ovule comparisons were performed individually with 0 DPA as a control point for comparison (Figure 3.1A).

Laser capture microdissection (LCM)

Cotton ovules collected at -2 DPA, 0 DPA, and 2 DPA were subjected to tissue fixation for laser capture microdissection whereas ovules at 7 DPA were processed by the method as described in Chapter II. Freshly collected ovules were fixed in a fixative (3:1 ethanol :

acetic acid) for 10 min and vacuum infiltrated on ice for 20 min. The infiltrated tissues were incubated at 4 ºC for 1 hour with rotation. The vacuum infiltration and incubation process were repeated twice with fresh fixative. Tissues infiltrated with 10 ml of 10% sucrose for 15 min, rotated overnight at 4 ºC and repeated with 15% sucrose for cryoprotection. The fixed ovules were embedded with Tissue-Tek® Optimal Cutting Temperature (O.C.T.) (Sakura Finetek U.S.A., Torrance, CA) in cryo-mold. The embedded ovules were frozen immediately in liquid nitrogen and stored at -80 ºC. Cryosectioning was performed with a Leica Cryostat (Leica Microsystems, Bannockburn, IL) in the microscopy facility at the University of Texas at Austin. The block was equilibrated at -20 ºC for 1 hour and cryosectioned at 10 µm. The slides with cryosectioned ovules were dehydrated with a series of ethanol (70%, 95%, and 100%) form 2 min each on ice and transferred to histoclear (National Diagnostics, Atlanta, GA).

The PALM laser capture system (P.A.L.M. Microlaser Technologies AG Inc., Bernried, Germany) was used for laser capture microdissection. Individual fiber initials (-2 DPA and 0 DPA) or epidermal cells (2 DPA) were catapulted (Figure 3.1B) and then 45 µl of RNALater (Ambion, Austin, TX) was added.

Microarray hybridization and statistical analysis

RNA of ovules at 7 DPA was extracted by the following procedure. Captured cells were homogenized in 500 µl of RNA extraction buffer (200 mM sodium borate decahydrate (pH 9), 30 mM ethylene glycol bis-N,N'-tetraacetic acid (EGTA), 1% (w/v) sodium

dodecyl sulfate (SDS), 1% (w/v) sodium deoxycholate, 2% (w/v) polyvinylpyrrolidone (PVP), 0.5% (v/v) Nonidet-40 (NP-40), 10 mM dithiothreitol (DTT)). After 250 µl of ethanol was added to the captured cells, the sample was loaded onto a Qiagen RNeasy mini column (Qiagen, Valencia, CA) and washed as per the manufacturer's recommendation.

For microarray hybridization, RNA was amplified with a Amino Allyl MessageAmp™ aRNA amplification Kit (Ambion, Austin, TX). Two rounds of amplication were performed to get sufficient quantities of aRNA. Amplified RNA was coupled with Cy3 or Cy5 dyes (Amersham Biosciences, Piscataway, NJ) and purified using Qiagen RNeasy mini column (Qiagen, Germantown, MD). About 1 µg of fragmented Cy3- and Cy5-labeled aRNA probes was used for hybridization. In each experiment, two dye-swaps with two biological replications were done in a total of eight slides. Hybridization and washing were performed as described in Chapter II.

After the data were processed using logarithm ratios of green and red hybridization signals, a robust and locally weighted linear regression (lowess) (Cleveland, 1979) was used to remove non-linear components (e.g. dye and pin effects) (Quackenbush, 2002). GeneSpring (Agilent Technologies, Foster City, CA) software was used to perform t-tests on the gene expression data and statistical significance was accepted at $p \leq 0.05$. Genes showing a significant p value were further analyzed by fold-expression. For clustering analysis, K-means clustering was applied with Euclidean distance measurement.

Quantitative RT-PCR (qRT-PCR) analysis

For microarray data verification, the same RNA used in the microarray analysis was used for qRT-PCR analysis. Two of the non-fiber tissues (leaves and petal) were added in the analysis to find the genes with fiber-specific expression. For the transcript amplification, gene-specific primers were designed using Primer Express version 2.0 software. The qRT-PCR reaction was carried out in a final volume of 20 µl containing 10 µl SYBR Green PCR master mix (Applied Biosystems, Foster City, CA), 1 µM forward and reverse primers, and 0.1 µM cDNA probe in a ABI7500 Real-Time PCR system (Applied Biosystems, Foster City, CA). Cotton *HISTONE3* (AF024716) was used to normalize the amount of gene-specific RT-PCR products (Wang et al., 2004). All reactions were performed in three replications using a dissociation curve to control the absence of primer dimers in the reactions. The amplification data was analyzed using ABI7500 SDS software (version 1.2.2) and the fold changes were calculated using the standard in each reaction.

Cloning and transformation

Full length cDNA of RDL gene (AY072821) was amplified from epidermis of 0 DPA ovules using gene specific primers with restriction enzyme sites (forward primer with *BamH*I site; 5'- CGCGGATCCGCGGGGAATTAGTCACTCCTGTTCTAG-3', reverse primer with *Not*I site; 5'- ATAAGAATGCGGCCGCTAAACTATTTACTTAGG

GACCCAAACAATGTGA). The amplified fragment was inserted into pGEM T-easy vector (Promega, Madison, WI) and sequenced. The pGEM T-easy vector containing *GhRDL* was digested with *BamH*I and *Not*I for cloning into a binary vector. The *BamH*I/*Not*I insert was then cloned into *BamH*I and *Not*I-cut 35SpBARN plant expression vector, making *35S:GhRDL* (LeClere and Bartel, 2001).

Binary construct was introduced into *Agrobacterium tumefaciens* strain GV3101 by electroporation and subsequently transferred into wild type (Col-0) by the floral dip method (Clough and Bent, 1998). As a control, pBARN vector without any transgene was transformed with the same procedure. For GUS assay, pRDL was amplified using gene specific primers with restriction enzyme sites (forward primer with *Hind*III site: 5'- CCCAAGCTTGGGGATAAC CTCTTAATCCACTAATGA- 3', reverse primer with SmaI site: 5'- TCCCCCGGGGG AGTGCTTTTCTCCTCCATCAGTT- 3') and inserted into pBI101. The transformation was performed as described above.

DNA extraction and genotyping

Genomic DNA extraction was performed as previously described (Chen et al., 1998; Madlung et al., 2002). Genotyping was performed using gene specific primers used for cloning. To ensure the size of the transgene, the insert was amplified from the binary vector constructed for transformation.

Histochemical GUS statining

Histochemical GUS assay was performed as described (Jefferson et al., 1987). The seeds were incubated in the staining solution for two days at 37ºC and stored in 70% ethanol before observation.

Scanning Electron Microscopy (SEM)

SEM was performed using a modified protocol (Murai et al., 2002; Tian et al., 2003). In brief, *Arabidopsis* seeds and leaves from T2 generations were fixed in a solution containing 3% each of formaldehyde and glutaraldehyde in 0.1 M sodium cacodylate buffer (pH = 7.4) and rinsed in 0.2 M sodium cacodylate buffer (pH = 7.4) three times. The ovules were washed in an ethanol series from 10 to 70% every 15 min. The concentration of the ethanol was increased to 100% within 18 h to dehydrate the samples for SEM analysis. The specimens were prepared by critical-point, dried with CO_2 at 1,400 and 1,800 psi, consecutively, and mounted by conductive gold paint and sputter coating with Cressington 208 coating. The samples were then scanned and analyzed using a Zeiss Supra 40 VP SEM located in the microscopy facility at the University of Texas at Austin with an accelerating voltage of 5 kV and a working distance of 39 mm. Images were scanned and stored as TIFF files.

CHAPTER IV

SUMMARY AND CONCLUSION

Cotton fibers are seed trichomes derived from individual cells of epidermal layer of cotton seed. Protodermal cells committed to cotton fiber undergo four overlapping developmental processes: fiber initiation, elongation, cellulose biosyhthesis, and maturation (Basra and Malik, 1984; Tiwari and Wilkins, 1995; Kim and Triplett, 2001). In spite of an economical importance of cotton in the world, the molecular basis of cell differentiation and fiber development is poorly understood. This study helps understand molecular and cellular events of cotton fiber development.

By comparing naked seed mutants (*N1N1* and *n2n2*) with its isogenic lines of cotton (*Gossypium hirsutum*, TM-1), we found that the dominant mutation (*N1N1*) that only produces very few lint fibers and no fuzz fibers delayed the process of fiber cell formation. This dominant mutant (*N1N1*) also showed reduced the total number of fiber cell initials, resulting in sparsely distributed short fibers (Figure 2.1 and Figure 2.2). Different gene expression profiles were observed between the dominant mutant (*N1N1*) and wild type cotton (TM-1) among four tissues using a pilot set of 70-mer oligonucleotide microarray that contained 1,536 features designed from a subset of cotton fiber ESTs (Table 2.1). Gene expression studies with fibers and non-fiber tissues suggest

with 20–50% fiber ESTs were expressed in both fibers (ovules at different developmental stages) and non-fiber tissues (leaf and stem). *N1N1* had both negative and positive effects on gene regulation associated with fiber development. Statistical analysis of microarray data and quantitative RT-PCR found 23 "fiber-associated" genes (Table 2.5 and Figure 2.5). These fiber associated genes played a role in the temporal regulation of genes involved in transcriptional and translational regulation, signal transduction, and cell differentiation during early stages of fiber development. They were concertedly down-regulated in the *N1N1* mutant leading to a defective process of fiber development.

To get a large view of fiber initiation, we developed an ovule EST library from *Gossypium hirs*utum L. TM-1 immature ovules (GH_TMO). The ESTs in this library represented approximately 15% of the unique sequences in the cotton EST collection and showed a significant increase in the percentage of genes encoding putative transcription factors and predicted proteins involved in hormone signaling pathways (Yang et al., 2006b). A new cotton oligonucleotide microarray was constructed from this GH_TMO EST library, a set from Jonathan Wendel's lab at Iowa State University, and the pilot set of oligos used for previous study.

Global gene expression studies were performed with the microdissected fiber initials (or epidermis) and the inner ovules using a large set of cotton microarray. It is necessary to start with fiber cells separated from the ovules to investigate fiber preferentially expressed genes because cotton fibers are derived from epidermal layers. Laser capture microdissection and aRNA amplification have been applied to various

studies with populations of specific cell types (Nakazono et al., 2003; Park et al., 2004; Casson et al., 2005). This method allowed us to collect fiber initials (0 DPA and 2 DPA) or epidermal layers (-2 DPA) from whole ovule tissues and get enough amount of probes for microarray hybridization. The gene expression profile of fiber initials was compared with gene expression of inner ovules at each stage (-2 DPA, 0 DPA, 2 DPA, and 7 DPA). In addition, differences of 4 different developmental stages were compared in inner ovules or epidermis with 0 DPA tissue as a control point (Figure 3.1).

Several hundreds of genes were identified in each experiment (Table 3.1). Up-regulated genes in the epidermis during fiber initiation stage (-2 DPA, 0 DPA, and 2 DPA) contain fiber proteins, several MYB transcription factors, and hormonal regulators. Enrichment of cotton *GL2* ortholog transcripts in the fiber cells of 7 DPA ovules (Table 3.2) confirmed the close relationship between cotton fiber development and *Arbidopsis* trichome development (Lee et al., 2007). In addition, the genes encoding *TESTAPARENT TESTA 2* (*TT2*) that interacts with *TTG1* for seed development were up-regulated in the fiber initials at 2 DPA (Table 3.2).

In each developmental stage, different sets of gene categories in molecular function or biological processes were over- or under-represented, suggesting temporal regulation of genes during fiber development (Figure 3.4). Indeed, clustering analysis of differentially expressed genes showed temporal regulation of gene expression (Figure 3.5).

One of the possible "fiber associated genes" found in microarray analyses, *RD22 like* gene (*GhRDL*), showed transcriptome abundance in the epidermis of cotton ovules during fiber initiation and its expression level was confirmed by quantitative RT-PCR (Figure 3.3). *RDL* in *Gossypium arboretum* (*GaRDL1*) is known to give trichome-specific expression in *Arabidopsis thaliana* (Wang et al., 2004). The *Arabidopsis* trichome system was used to investigate the function of *GhRDL* because leaf trichome development in *Arabidopsis* and fiber development in cotton share many similarities in terms of cell differentiation.

For overexpression and complementation study, *35S:GhRDL* was constructed in a binary vector pBARN and transformed into *Arabidopsis thaliana* columia-0 and glabrous mutant *gl1* (CS1644). It was expected to observe complementation of leaf trichomes in the mutant background or over-produced leaf trichomes in the wild type background, assuming that *GhRDL* can play an important role in the trichome development pathway. However, *35S*:*GhRDL* failed to complement glabrous phenotype in *gl1*. It is possible that *GhRDL* and *GL1* are not in the same pathway or cDNA of *GhRDL* is not sufficient to complement the phenotype. Although complementation test in *gl1* didn't work out, in the transgenic plants harboring *35S:GhRDL*, the number of trichomes on leaves was significantly higher and seed hairs (or seed trichomes) were also induced in overexpression lines (Figure 3.8 and 10). It was an interesting result because cotton fibers are seed trichomes derived from epidermal layers of cotton ovules. Furthermore, the promoter of *GhRDL* showed the activity in the seeds (Figure 3.11) indicating *GhRDL* is

acting particularly in seeds. Taken together, the data suggest that *RDL* in *Gossypium hirsutm* is preferentially expressed in fiber cells and a key player in cell differentiation and fiber development.

References

Abe, M., Takahashi, T., and Komeda, Y. (1999). Cloning and characterization of an L1 layer-specific gene in *Arabidopsis thalian*. Plant Cell **45,** 571-580.

Abe, M., Takahashi, T., and Komeda, Y. (2001). Identification of a cis-regulatory element for L1 layer-specific gene expression, which is targeted by an L1-specific homeodomain protein. Plant J **26,** 487-494.

Adams, K.L., Croon, R., Percifield, R., and Wendel, J.F. (2003). Genes duplicated by polyploidy show unequal contributions to transcriptome and organ-specific reciprocal silencing. Proc Natl Acad Sci **100,** 4649-4654.

Angeles, G., Berrio-Sierra, J., Joseleau J-P., Lorimier, P., Lefevre A., and K., R. (2006). Preparative laser capture microdissection and single-pot cell wall material preparation: a novel method for tissue-specific analysis. Planta **224,** 228-232.

Applequist, W.L., Cronn, R., and Wendel, J.F. (2001). Comparative development of fiber in wild and cultivated cotton. Evolution and Development **3,** 3-17.

Arpat, A.B., Waugh, M., Sullivan, J.P., Gonzales, M., Frisch, D., Main, D., Wood, T., Leslie, A., Wing, R.A., and Wilkins, T.A. (2004). Functional genomics of cell elongation in developing cotton fibers. Plant Mol Biol **54,** 911-929.

Basra, A.S., and Malik, C.P. (1984). Development of the cotton fiber. International Review of Cytology **89,** 65-113.

Basra, A.S., and Saha, S. (1999). Growth regulation of cotton fibers. In Cotton fibers, A.S. Basra, ed (Binghamton, NY: The Haworth Press), pp. 47-63.

Baudry, A., Heim, M.A., Dubreucq, B., Caboche, M., Weisshaar, B., and Lepiniec, L. (2004). TT2, TT8, and TTG1 synergistically specify the expression of *BANYULS* and proanthocyanidin biosynthesis in *Arabidopsis thaliana.* Plant J **39,** 366-380.

Beasley, C.A. (1971). In vitro culture of fertilized cotton ovules. BioScience **21,** 906-907.

Beasley, C.A., and Ting, I.P. (1974). Effects of plant growth substances on in vitro fiber development from unfertilized cotton ovules. Am J Bot **61,** 188-194.

Benedict, C.R. (1984). Physiology. In Cotton., R. J. Kohel and C.F. Lewis, eds (Madison, WI: American Society of Agronomy.

Benedict, C.R., Smith, R.H., and Kohel, R.J. (1973). Incorporation of ^{14}C-photosynthate into developing cotton bolls, *Gossypium hirsutum* L. Crop Science **13,** 88-91.

Berlin, J.D. (1986). The outer epidermis of the cotton seed. In Cotton physiology, J.R. Mauney and J.M. Stewart, eds (Memphis, TN: The Cotton Foundation), pp. 375-414.

Black, M.A., and Doerge, R.W. (2002). Calculation of the minimum number of replicate spots required for detection of significant gene expression fold change in microarray experiment. Bioinformatics **18,** 1609–1616.

Blanc, G., and Wolfe, K.H. (2004a). Functional divergence of duplicated genes formed by polyploidy during *Arabidopsis* evolution. Plant Cell **16,** 1679–1691.

Blanc, G., and Wolfe, K.H. (2004b). Widespread paleopolyploidy in model plant species inferred from age distributions of duplicate genes. Plant Cell **16,** 1667–1678.

Cai, S., and Lashbrook, C.C. (2006). Laser capture microdissection of plant cells from tape-transferred paraffin sections promotes recovery of structurally intact RNA for global gene profiling. Plant J **48**.

Casson, S., Spencer, M., Walker, K., and Linsey, K. (2005). Laser capture microdissection for the analysis of gene expression during embryogenesis for *Arabidopsis*. Plant J **42,** 111-123.

Chen, Z.J., Comai, L., and PIKAARD, C.S. (1998). Gene dosage and stochastic effects determine the severity and direction of uniparental ribosomal RNA gene silencing (nucleolar dominance) in *Arabidopsis* allopolyploids. Proc Natl Acad Sci **95,** 14891–14896.

Chen, Z.J., Wang, J.L., Tian, L., Lee, H.S., Wang, J.Y.J., Chen, M., Lee, J.S.J., Josefsson, C., Madlung, A., Watson, B., Lippman, Z., Vaughn, M., Pires, J.C., Colot, V., Doerge, R.W., Martienssen, R.A., Comai, L., and Osborn, T.C. (2004). The development of an *Arabidopsis* model system for genome-wide analysis of polyploidy effects. Bio J Linn Soc **82,** 689-700.

Cherry, J.P., Katterman, F.R.H., and Endrizzi, J.E. (1972). Seed esterases, leucine aminopepridases, and catalases of species of the genus *Gossypium.* Theor Appl Genet **42,** 218-226.

Chou, H.-H., Hsia, A.-P., Mooney, D.L., and Schnable, P.S. (2004). Picky: oligo microarray design for large genomes. . Bioinformatics **20,** 2893-2902.

Cleveland, W.S. (1979). Robust locally weighted regression and smoothing scatterplots. J Am Stat Asso **74,** 829–836

Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana.* Plant J **16,** 735-743.

Cronn, R.C., Small, R.L., and Wendel, J.F. (1999). Duplicated genes evolve independently after polyploid formation in cotton. Proc Natl Acad Sci **96,** 14406-14411.

Day, R.C., McNoe, L.A., and Macknight, R.C. (2007). Transcript analysis of laser microdissected plant cells. Physiologia Plantarum **129,** 267-282.

Desai, A., Chee, P.W., Rong, J., May, O.L., and Paterson, A.H. (2006). Chromosome structural changes in diploid and tetraploid A genomes of *Gossypium*. Genome **49,** 336-345.

Dixon, D.C., Seagull, R.W., and Triplett, B.A. (1994). Changes in the Accumulation of [alpha]- and [beta]-Tubulin Isotypes during Cotton Fiber Development. Plant Physiol **105,** 1347-1353.

Endrizzi, J.E., Turcotte, E.L., and Kohel, R.J. (1984). Qualitative genetics, cytology and cytogenetics. In Agronomy: Cotton, R.J. Kohel and D.F. Lewis, eds (Madison, WI: American Society of Agronomy Inc), pp. 59-80.

Endrizzi, J.E., Turcotte, E.L., and RJ, K. (1985). Genetics, cytology, and evolution of *Gossypium.* Adv Genet **23,** 271-375.

Espina, V., Wufkuhle, J.D., Calvert, V.S., VanMeter A., Zhou, W., Coukos, G., Geho, D.H., Petricoin III, E.F., and Liotta, L.A. (2006). Laser-capture microdissection. Nature Protocols **1 (2),** 586-603.

Fojtova, M., Kovarik, A., Votruba, I., and Holy, A. (1998). Evaluation of the impact of S-adenosylhomocysteine metabolic pools on cytosine methylation of the tobacco genome. Eur J Biochem **252,** 347-352.

Fryxell, P.A. (1963). Morphology of the base of seed hairs of *Gossypium.* Botanical Gazette **124,** 196-199.

Fryxell, P.A. (1979). The natural history of the cotton tribe. (College Station, TX: Texas A&M University).

Fryxell, P.A. (1992). A revised taxomomic interpretation of *Gossypium* L. (Malvaceae). Rheedea **2,** 108-165.

Gallardo, K., Job, C., Groot, S.P., Puype, M., Demol, H., Vandekerckhove, J., and Job, D. (2002). Proteomics of Arabidopsis seed germination. A comparative study of wild-type and gibberellin-deficient seeds. Plant Physiol **129,** 823-837.

Gao, P., Zhao, P.-M., Wanga, J., Wang, H.-Y., Wu X-M., and Xia, G.-X. (2007). Identification of genes preferentially expressed in cotton fibers: A possible role of calcium signaling in cotton fiber elongation Plant Science **173,** 61-69.

Gialvalis, S., and Seagull, R.W. (2001). Plant hormones alter fiber initiation in unfertilized, cultured ovules of *Gossypium hirsutum.* J Cotton Sci **5,** 252-258.

Guinn, G., and Brummett, D.L. (1988). Changes in abscisic acid and indoleacetic acid before and after anthesis relative to changes in abscission rates of cotton fruiting forms. Plant Physiol **87,** 629-631.

Hendrix, B., and Stewart, J.M. (2005). Estimation of the nuclear DNA content of *Gossypium* species. Annal of Bot **95,** 789-797.

Hendrix, D.L. (1990). Carbohydrates and carbohydrate enzymes in developing cotton ovules. Physiologia Plantarum **78,** 85-92.

Hochberg, Y., and Tamhane, A.C. (1987). Multiple comparison procedures. (New York: Wiley).

Hülskamp, M. (2004). Plant trichomes: A model for cell differentiation. Nature Rev Mol Cell Biol **5,** 471-480.

Hülskamp, M., and Schnittger, A.T. (1998). Spatial regulation of trichome formation in *Arabidopsis thaliana.* Seminars in Cell and Developmental Biology **9,** 213-220.

Hülskamp, M., Misera, S., and Jurgens, G. (1994). Genetic dissection of trichome cell development in *Arabidopsis.* Cell **76,** 555-566.

Humphries, J.A., Walker, A.R., Timmis, J.N., and Orford, S.J. (2005). Two WD-repeat genes from cotton are functional homologues of the *Arabidopsis thaliana TRANSPARENT TESTA GLABRA1 (TTG1)* gene. Plant Mol Biol **57,** 67-81.

Ishida, T., Hattori, S., Sano, R., Inoue, K., Shirano, Y., Hayashi, H., Shibata, D., Sato, S., Kato, T., Tabata, S., and Okada, K., Wada, Takuji. . (2007). *Arabidopsis TRANSPARENT TESTA GLABRA2* is directly regulated by R2R3 MYB transcription factors and is involved in regulation of *GLABRA2* transcription in epidermal differentiation. Plant Cell **19,** 2531-2543.

Jefferson, R.A., Kavanagh, T.A., and Bevan, M.W. (1987). GUS fusions: β-Glucuronidase as a seneitive and versatile gene fusion marker in higher plants. Embo J **6,** 3901-3907.

Jenson, S.D., Robetorye, R.S., Bohling, S.D., Schumacher, J.A., Morgan, J.W., Lim, M.S., and Elenitoba-Johnson, K.S.J. (2003). Validation of cDNA microarray gene expression data obtained from linearly amplified RNA. J Clin Pathol: Mol Pathol **56,** 307-312.

Jernstedt, J.A., Sugimoto, K., Lu, P., and Wilkins, T.A. (1993). Microtubule involvement in cotton fiber initiation. Am J Bot **80,** 28.

Ji, S., Lu, Y., Li, J., Wei, G., Liang, X., and Zhu, Y. (2002). A beta-tubulin-like cDNA expressed specifically in elongating cotton fibers induces longitudinal growth of fission yeast. Biochem Biophys Res Commun **296,** 1245-1250.

Ji, S.J., Lu, Y.C., Feng, J.X., Wei, G., Li, J., Shi, Y.H., Fu, Q., Liu, D., Luo, J.C., and Zhu, Y.X. (2003). Isolation and analyses of genes preferentially expressed during early cotton fiber development by subtractive PCR and cDNA array. Nucleic Acids Res **31,** 2534-2543.

Jin, L., Thompson, C.A., Zian, X., Kuecker, S.J., Kuling, E., and Lloyd, R. (1999). Analysis of anterior pituitary hormone mRNA expression in immunophenotypically characterized single cells after laser capture microdissection. Lab Investig **79,** 511-512.

John, M.E. (1996). Structural characterization of genes corresponding to cotton fiber mRNA, E6: reducaed E6 protein in transgenic plants by antisense gene. Plant Mol Biol **30,** 297-306.

John, M.E., and Crow, L.J. (1992). Gene expression in cotton (*Gossypium hirsutum* L.) fiber: Cloning of the mRNAs. Proc Natl Acad Sci **89,** 5769-5773.

Joshi, P.A., Stewart, J.M., and Graham, E.T. (1985). Localization of β-glycerophosphatase activity in cotton fiber during differentiation. Protoplasma **125,** 75-85.

Kane, M.D., Jatkoe, T.A., Stumpf, C.R., Lu, J., Thomas, J.D., and Madore, S.J. (2000). Assessment of the sensitivity and specificity of oligonucleotide (50 mer) microarrays. Nucleic Acids Res **28,** 4552-4557.

Kerk, N.M., Ceserani, T., Tausta, S.L., Sussex, I.M., and Nelson, T.M. (2003). Lase capture microdissection of cells from plant tissues. Plant Physiol **132,** 27-35.

Kerr, M.K., and Churchill, G.A. (2001). Experimental design for gene expression microarrays. Bioinformatics **2,** 183–201.

Kim, H.J., and Triplett, B.A. (2001). Cotton fiber growth in planta and in vitro. Models for plant cell elongation and cell wall biogenesis. Plant Physiol **127,** 1361-1366.

Kim, H.J., Pesacreta, T.C., and Triplett, B.A. (2004c). Cotton-fiber germin-like protein. II: Immunolocalization, purification, and functional analysis. Planta **218,** 525-535.

Kohel, R.J. (1973). Genetic nomenclature in cotton. J Hered **64**.

Kosmidou-Dimitropoulou, K. (1986). Hormonal influences on fiber development. In Cotton physiology, J.R. Mauney and J.M. Stewart, eds (Memphis, TN: The Cotton Foundation), pp. 361-373.

Lamblin, F., Saladin, G., Dehorter, B., Cronier, D., Grenier, E., Lacoux, J., Bruyant, P., Lainé, E., Chabbert, E., Girault, F., Monties, B., Morvan, C., David, H., and David, A. (2001). Overexpression of a heterologous sam gene encoding S-adenosylmethionine synthetase in flax (*Linum usitatissimum*) cells: Consequences on methylation of lignin precursors and pectins. Physiologia Plantarum **112,** 223-232.

Larkin, J.C., Oppenheimer, D.G., Pollock, S., and Mark, M.D. (1993). *Arabidopsis GLABROUS1* gene requires downstream sequences for function. Plant Cell **5,** 1739-1748.

LeClere, S., and Bartel, B. (2001). A library of *Arabidopsis* 35S-cDNA lines for identifying novel mutants. Plant Mol Biol **46,** 695-703.

Lee, H.S., Wang, J.L., Tian, L., Jiang, H.M., Black, M.A., Madlung, A., Watson, B., Lukens, L., Pires, J.C., Wang, J.J., Comai, L., Osborn, T.C., Doerge, R.W., and Chen, Z.J. (2004). Sensitivity of 70-mer oligonucleotides and cDNAs for microarray analysis of gene expression in *Arabidopsis* and its related species. Plant Biotech J **2,** 45-57.

Lee, J.J., Woodward, A.W., and Chen, Z.J. (2007). Gene expression changes and early events in cotton fibre development. Annal of Bot **100,** 1391-1401.

Lee, J.J., Hassan, O.S., Gao, W., Wei, N.E., Kohel, R.J., Chen, X.Y., Payton, P., Sze, S.H., Stelly, D.M., and Chen, Z.J. (2006). Developmental and gene expression analyses of a cotton naked seed mutant. Planta **223,** 418-432.

Li, C.H., Zhu, Y.Q., Meng, Y.L., Wang, J.W., Xu, K.X., Zhang, T.Z., and Chen, X.Y. (2002a). Isolation of genes preferentially expressed in cotton fibers by cDNA filter arrays and RT-PCR. Plant Science **163,** 1113-1120.

Li, F., and Stormo, G.D. (2001). Selection of optimal DNA oligos for gene expression arrays. Bioinformatics **17,** 1067-1076.

Li, X.B., Cai, L., Cheng, N.H., and Liu, J.W. (2002b). Molecular characterization of the cotton GhTUB1 gene that is preferentially expressed in fiber. Plant Physiol **130,** 666-674.

Li, X.B., Fan, X.P., Wang, X.L., Cai, L., and Yang, W.C. (2005). The cotton *ACTIN1* gene is functionally expressed in fibers and participates in fiber elongation. Plant Cell **17,** 859-875.

Loguercio, L.L., Zhangg, J.Q., and Wilkins, T.A. (1999). Differential regulation of six novel *MYB*-domain genes defines two distinct expression patterns in allotetraploid cotton (*Gossypium hirsutum* L.). Mol Genet Genomics **261,** 660-671.

Lolle, S.J., Cheung, A.Y., and Sussex, I.M. (1992). Fiddlehead: an *Arabidopsis* mutant constitutively expressing an organ fusion program that involves interactions between epidermal cells. Developmental Biol **152,** 383-392.

Luschnig, C. (2002). Auxin transport: ABC proteins join the club. Trends in Plant Science **7,** 329-332.

Madlung, A., Ricardo, W.M., Watson, B., Reynolds, S.H., Davison, J., and Comai, L. (2002). Remodeling of DNA methylation and phenotypic and transcriptional changes in synthetic *Arabidopsis* allotetraploids. Plant Physiol **129,** 733-746.

Marks, M.D. (1997). Molecular genetic analysis of trichome development in *Arabidopsis.* Annu Rev Plant Physiol Plant Mol Biol **48,** 137-163.

Matsunaga, S., Schutze, K., Donnison I.S., Grant, S.R., Kuroiwa, T., and Kawano, S. (1999). Single pollen typing combined with laser-mediated manipulation. Plant J **20,** 371-378.

Mauney, J.R. (1984). Anatomy and morphology of cultivated cotton. In Cotton, R. Kohel, Lewis, C.F., ed (Madison, WI: American Society of Agronomy.

May, O.L., and Lege, K.E. (1999). Development of the world cotton industry. In Cotton: Origin, History, Technology, and Production, W.C. Smith and J.T. Cothren, eds (New York: John Wiley & Sons), pp. 65.

Meinert, M.C., and Delmer, D.P. (1977). Changes in biochemical composition of cell-wall of cotton fiber during development. Plant Physiol **59,** 1088–1097.

Murai, K., Takumi, S., Koga, H., and Ogihara, Y. (2002). Pistillody, homoeotic transformation of stamens into pistil-like structures, caused by nuclear-cytoplasm interaction in wheat. Plant J **29**.

Nakazono, M., Qiu, F., Borsuk, L.A., and Schnable, P.S. (2003). Laser-capture microdissection, a tool for the golobal analysis of gene expression in specific plant cell types: identification of gene expressed differentilly in epidermal cells or vascular tissues of maize. Plant Cell **15,** 583-596.

Nolte, K.D., Hendrix, D.L., Radin, J.W., and Koch, K.E. (1995). Sucrose synthase localization during Initiation of seed development and trichome differentiation in cotton ovules. Plant Physiol **109,** 1285-1293.

Ohyama, H., Zhang, X., Kohno, Y., Alevizos, I., Posner, M., Wong, D.T., and Todd, R. (2000). Laser capture microdissection-generated target sample for high-density oligonucleotide array hybridization. Biotechniques **29,** 530-536.

Park, J.Y., Kim, S.Y., Lee, J.H., Song, J., Noh, J.H., Lee, S.H., Park, W.S., Yoo, N.J., Lee, J.Y., and Nam, S.W. (2004). Amplification of amplified RNA and evaluation of cRNA targets for spotted-oligonucleotide microarray. Biochem Biophys Res Commun **325,** 1346-1352.

Payne, C., Zhang, F., and Lloyd, A. (2000). GL3 encodes a bHLH protein that regulate trichome development in Arabidopsis through interaction with GL1 and TTG1. Genetics **156,** 1349-1362.

Payne, T., Clement, J., Arnold, D., and Lloyd, A. (1999). Heterologous myb genes distinct from *GL1* enhance trichome production when overexpressed in *Nicotiana tabacum.* Development **126,** 671-682.

Percival, A.E., Wendel JF, and Stewart, J.M. (1999). Taxonomy and Germplasm Resources. In Cotton: Origin, History, Technology, and Production, W. Smith and J.T. Cothren, eds (New York: John Wiley & Sons), pp. 33-63.

Pesch, M., and Hülskamp, M. (2004). Creating a two-dimentional pattern *de novo* during *Arabidopsis* trichome and root hair initiation. Curr Opin Genet Devel **14,** 422-427.

Pruitt, R.E., Vielle-Calzada, J.P., Ploense, S.E., Grossniklaus, U., and Lolle, S.J. (2000). *FIDDLEHEAD*, a gene required to suppress epidermal cell interactions in *Arabidopsis*, encodes a putative lipid biosynthetic enzyme. Proc Natl Acad Sci **97,** 1311-1316.

Quackenbush, J. (2002). Microarray data normalization and transformation. Nature Genetics **32,** 496-501.

Ramsey, J.C., and Berlin, J.D. (1976). Ultrastructure of early stages of cotton fiber differentiation. Botanical Gazette **137,** 11-19.

Rerie, W.G., Feldmann, K.A., and Marks, M.D. (1994). The *GLABRA2* gene encodes a homeo domain protein required for normal trichome development in *Arabidopsis.* Genes Devel **8,** 1388-1399.

Rook, M.S., Delach, S.M., Deyneko, G., Worlock, A., and Wolfe, J.L. (2004). Whole genome amplification of DNA from laser capture-microdissected tissue for high-throughput single nucleotide polymorphism and short tandem repeat genotyping. Am J Pathol **164,** 23-33.

Rouillard, J.M., Herbert, C.J., and Zuker, M. (2002). OligoArray: genome-wide oligonucleotide design for microarrays. . Bioinformatics **18,** 486-487.

Ruan, Y.L., and Chourey, P.S. (1998). A Fiberless seed mutation in cotton is associated with lack of fiber cell initiation in ovule epidermis and alterations in sucrose synthase expression and carbon partitioning in developing seeds. Plant Physiol **118,** 399-406.

Ruan, Y.L., Llewellyn, D., and Furbank, R. (2001). The control of single-celled cotton fiber elongation by developmentally reversible gating of plasmodesmata and coordinated expression of sucrose and K+ transporters and expansin. Plant Cell **13,** 47-60.

Ruan, Y.L., Llewellyn, D.J., and Furbank, R.T. (2003). Suppression of sucrose synthase gene expression represses cotton fiber cell initiation, elongation, and seed development. Plant Cell **15,** 952-964.

Ryser, U. (1985). Cell wall biosynthesis in differentiating cotton fibers. Eur J Cell Biol **39,** 236-256.

Ryser, U. (1999). Cotton fiber initiation and histodifferentiation. In Cotton fibers, A.S. Basra, ed (Binghamton, NY: The Haworth Press), pp. 1-45.

Schellmann, S., and Hülskamp, M. (2005). Epidermal differentiation: Trichomes in *Arabidopsis* as a model system. Intl J Devel Biol **49,** 579-584.

Schellmann, S., Schnittger, A., Kirik, V., Wada, T., Okada, K., Beermann, A., Thumfahrt, J., Jürgens, G., and Hülskamp, M. (2002). *TRIPTYCHON* and *CAPRICE* mediate lateral inhibition during trichome and root hair patterning in *Arabidopsis*. Embo J **21,** 5036-5046.

Schubert, A.M., Benedict, C.R., Berlin, J.D., and Kohel, R.J. (1973). Cotton fiber development - Kinetics of cell elongation and secondary wall thickening. Crop Science **13,** 704-709.

Seagull, R. (1990). The effects of microtubule and microfilament disrupting agents on cytoskeketal arrays and wall deposition in developing cotton fibers. Protoplasma **159,** 44-59.

Seagull, R. (1992a). A quantitative electron microscopic study of changes in microtubule arrays and wall microfibril orientation during in vitro cotton fiber development. J Cell Sci **101,** 561-577.

Seagull, R. (1992b). Cytoskeletal involvement in microfibril orgnization during cotton fiber development. In Preceedings from Cotton Fiber Cellulose:Structure, Function, and Utilization Conference (Memphis, TN: National Cotton Council), pp. 171-192.

Serna, L., and Martin, C. (2006). Trichomes: different regulatory networks lead to convergent structures. Trends Plant Sci **11,** 274-280.

Sharma, P., Singh, B., Malik, C.P., and Bajaj, K.L. (1995). Effect fo a plant growth retardant (E)-1-(4-chlorophenyl)-4,4-dimethyl-2-(1,2,4,-triazol-1-yl)-1-penten-3-ol (S-3307) on the fiber initiation and enzyme activities of cotton ovules (*Gossypium arboreum*). Phytomorphology **45,** 79-86.

Shi, Y.H., Zhu, S.W., Mao, X.Z., Feng, J.X., Qin, Y.M., Zhang, L., Cheng, J., Wei, L.P., Wang, Z.Y., and Zhu, Y.X. (2006). Transcriptome profiling, molecular biological, and physiological studies reveal a major role for ethylene in cotton fiber cell elongation. Plant Cell **18,** 651-664.

Smith, W.C., and Cothren, J.T. (1999). Cotton: Origin, history, technology, and production. (John Wiley & Sons, Inc.).

Suo, J., Liang, X., Pu, L., Zhang, Y., and Xue, Y. (2003). Identification of GhMYB109 encoding a R2R3 MYB transcription factor that expressed specifically in fiber initials and elongating fibers of cotton (*Gossypium hirsutum* L.). Biochimica et Biophysica Acta **1630,** 25-34

Szymanski, D.B., and Marks, M.D. (1998). *GLABROUS1* overexpression and *TRIPTYCHON* alter the cell cycle and trichome cell fate in *Arabidopsis*. Plant Cell **10,** 2047-2062.

Szymanski, D.B., Lloyd, A.M., and Marks, M.D. (2000). Progress in the molecular genetic analysis of trichome initiation and morphogenesis in *Arabidopsis.* Trends Plant Sci **5,** 214-219.

Taliercio, E.W., and Boykin, D. (2007). Analysis of gene expression in cotton fiber initials. BMC Plant Biology **7,** 22.

Taliercio, E.W., Hendrix, B., and Stewart, J.M. (2005). DNA content and expression of genes related to cell cycling in developing *Gossypium hirsutum* (*Malvaceae*) fibers. Am J Bot **92,** 1942-1947.

Tanaka, H., Masuta, C., Uehara, K., Kataoka, J., Koiwai, A., and Noma, M. (1997). Morphological changes and hypomethylation of DNA in transgenic tobacco expressing antisense RNA of the S-adenosyl-L-homocysteine hydrolase gene. Plant Mol Biol **35,** 981-986.

Tian, L., Wang, J., Fong, M.P., Chen, M., Cao, H., Gelvin, S.B., and Chen, Z.J. (2003). Genetic control of developmental changes induced by disruption of *Arabidopsis* histone deacetylase 1 (AtHD1) expression. Genetics **165,** 399-409.

Tian, L., Fong, M.P., Wang, J.J., Wei, N.E., Jiang, H., Doerge, R.W., and Chen, Z.J. (2005). Reversible histone acetylation and deacetylation mediate genome-wide, promoter-dependent and locus-specific changes in gene expression during plant development. Genetics **169,** 337-345.

Tiwari, S.C., and Wilkins, T.A. (1995). Cotton (*Gossypium hirsutum*) seed trichomes expand via diffuse growing mechanism. Can J Bot **73,** 746-757.

Turley, R.B., and Kloth , R.H. (2002). Identification of a Third Fuzzless Seed Locus in Upland Cotton (*Gossypium hirsutum* L.) J Hered **93,** 359-364.

Turley, R.B., and Kloth , R.H. (2008). The inheritance model for the fiberless trait in upland cotton (Gossypium hirsutum L.) line SL1-7-1: variation on a theme. Euphytica **in press**.

Udall, J.A., Flagel, L.E., Cheung, F., Woodward, A.W., Hovav, R., Rapp, R.A., Swanson, J.M., Lee, J.J., Gingle, A.R., Nettleton, D., Town, C.D., Chen, Z.J., and Wendel, J.F. (2007). Spotted cotton oligonucleotide microarrays for gene expression analysis. BMC Genomics **8,** 81.

Udall, J.A., Swanson, J.M., Haller, K., Rapp, R.A., Sparks, M.E., Hatfield, J., Yu, Y., Wu, Y., Dowd, C., Arpat, A.B., Sickler, B.A., Wilkins, T.A., Guo, J.Y., Chen, X.Y., Scheffler, J., Taliercio, E., Turley, R., McFadden, H., Payton, P., Klueva, N., Allen, R., Zhang, D., Haigler, C., Wilkerson, C., Suo, J., Schulze, S.R., Pierce, M.L., Essenberg, M., Kim, H., Llewellyn, D.J., Dennis, E.S., Kudrna, D., Wing, R., Paterson, A.H., Soderlund, C., and Wendel, J.F. (2006). A global assembly of cotton ESTs. . Genome Res **16,** 441-450.

Van't Hof, J. (1999). Increased nuclear DNA content in developing cotton fiber cells. Am J Bot **86,** 776–779.

Van't Hof, J., and Saha, S. (1997). Cotton fibers can undergo cell division. Am J Bot **84,** 1231-1235.

Wang, J., Lee, J.J., Tian, L., Lee, H.S., Chen, M., Rao, S., Wei, E.N., Doerge, R.W., Comai, L., and Chen, Z.J. (2005a). Methods for genome-wide analysis of gene expression changes in polyploids. Methods Enzymol **395,** 570-596.

Wang, S., Wang, J.W., Yu, N., Li, C.H., Luo, B., Gou, J.Y., Wang, L.J., and Chen, X.Y. (2004). Control of plant trichome development by a cotton fiber MYB gene. Plant Cell **16,** 2323-2334.

Wendel, J.F. (1989). New World tetraploid cottons Old World cytoplasm. Proc Natl Acad Sci **86,** 4132-4136.

Wendel, J.F., and Albert, V.A. (1992). Phylogenetics of the cotton genus (*Gossypium* L.): Character-state weighted parsimony analysis of chloroplast DNA restriction site data and its systematic and biogeographic implications. Systematic Botany **17,** 115-143.

Wendel, J.F., and Cronn, R.C. (2003). Polyploidy and the evolutionary history of cotton. Adv Agron **78,** 139-186.

Wendel, J.F., Schnabel, F.A., and Seelanan, T. (1995). An unusual ribosomal DNA sequence from *Gossypium gossypoides* reveals ancient, cryptic, intergenomic introgression. Mol Phylo Evol **4,** 298-313.

Wilkins, T.A., and Jernstedt, J.A. (1999). Molecular genetics of developing cotton fibers. In Cotton fibers, A.S. Basra, ed (Binghamton, NY: Haworth Press), pp. 231-267.

Wu, J., Steinebrunner, I., Sun, Y., Butterfield, T., Torres, J., Arnold, D., Gonzalez, A., Jacob, F., Reichler S., and Roux, S.J. (2007a). Apyrases (Nucleoside Triphosphate-Diphosphohydrolases) Play a Key Role in Growth Control in Arabidopsis. Plant Physiol **144,** 961-975.

Wu, Y., Machado, A.C., White, R.G., Llewellyn, D.J., and Dennis, E.S. (2006). Expression profiling identifies genes expressed early during lint fibre initiation in cotton. Plant Cell Physiol **47,** 107-127.

Wu, Y., Llewellyn, D., White, R., Ruggiero, K., Al-Ghazi, Y., and Dennis, E.S. (2007b). Laser caputre microdissection and cDNA microarrays used to generate gene expression profiles of the rapidly expanding fibre initial cells on the surface of cotton ovules. Planta **226,** 1475-1490.

Yamaguchi-Shinozaki, K., Shinozaki, K. (1993). The plant hormone abscisic acid mediates the drought-induced expression but not the seed-specific expression of *rd22*, a gene responsive to dehydration stress in *Arabidopsis thaliana*. Mol Genet Genomics **238,** 17-25.

Yang, J.H., Han, S.J., Yoon, E.K., and Lee, W.S. (2006a). Evidence of an auxin signal pathway, microRNA167-ARF8-GH3, and its response to exogenous auxin in cultured rice cells. Nucleic Acids Res **34,** 1892-1899.

Yang, S.S., Cheung, F., Lee, J.J., Ha, M., Wei, N.E., Sze, S.-H., Stelly, D.M., Thaxton, P., Triplett, B., Town, C.D., and Chen, Z.J. (2006b). Accumulation of genome-specific transcripts, transcription factors, and phytohormonal regulators during early stages of fiber cell development in allotetraploid cotton. Plant J **47,** 761-775.

Yephremov, A., WIsman, E., Huijser, P., Huijser C., Wellesen K., and Saedker, H. (1999). Characterization of the *FIDDLEHEAD* gene of *Arabidopsis* reveals a link between adhesion response and cell differentiation in the epidermis. Plant Cell **11,** 2187-2201.

Zhang, F., Gonzalez, A., Zhao, M., Payne C.T., and Lloyd, A. (2003). A network of redundant bHLH proteins functions in all TTG1-dependent pathways of Arabidopsis. Development **130,** 4859-4869.

Zhu, B., Xu, F., and Baba, Y. (2006). An evaluation of linear RNA amplification in cDNA microarray gene expression analysis. Mol Genet Metab **87,** 71-79.

Zhu, Y.Q., Xu, K.X., Luo, B., Wang, J.W., and Chen, X.Y. (2003). An ATP-binding cassette transporter *GhWBC1* from elongating cotton fibers. Plant Physiol **133,** 580-588.

VITA

Jinsuk Lee was born in Seoul, South Korea on Oct. 8th, 1978, the daughter of Seungkook Lee and Sungsoon Kim. She received the degree of Bachelor of Science from Seoul National University in Seoul, South Korea in 2000. She started the graduate studies at Texas A&M University in 2000 and received the degree of Master of Science in 2002.

ADDRESS 106-904 Korng-apt, Sukwan-dong, Sungbuk-gu,

Seoul, South Korea 136-762

PUBLICATIONS

Lee JJ, Woodward AW, Chen ZJ. Gene expression changes and early events in cotton fiber development. Annals of Botany. 2007. 100(7):1391-401

Udall JA, Flagel LE, Cheung F, Woodward AW, Hovav R, Rapp RA, Swanson JM, Lee JJ, Gingle AR, Nettleton D, Town CD, Chen ZJ, Wendel JF. Spotted cotton oligonucleotide microarrays for gene expression analysis. BMC Genomics.

2007. 8:81.

Samuel Yang S, Cheung F, Lee JJ, Ha M, Wei NE, Sze SH, Stelly DM, Thaxton P, Triplett B, Town CD, Jeffrey Chen Z. Accumulation of genome-specific transcripts, transcription factors and phytohormonal regulators during early stages of fiber cell development in allotetraploid cotton. Plant J. 2006. 47(5): 761-75.

Lee JJ, Hassan OS, Gao W, Wei NE, Kohel RJ, Chen XY, Payton P, Sze SH, Stelly DM, Chen ZJ. Developmental and gene expression analyses of a cotton naked seed mutant. Planta. 2006. 223(3):418-32.

Lee JJ, Crosby KM, Pike LM, Yoo KS, Leskovar DI. Impact of genetic and environmental variation on development of flavonoids and carotenoids in pepper (*Capsicum* spp.). Scientia Horticulturae. 2005. 106:341-352.

Wang J, Lee JJ, Tian L, Lee HS, Chen M, Rao S, Wei EN, Doerge RW, Comai L, Chen ZJ. Methods for genome-wide analysis of gene expression changes in polyploids. Methods Enzymol. 2005. 395:570-96.

Wang J, Tian L, Madlung A, Lee HS, Chen M, Lee JJ, Watson B, Kagochi T, Comai L, Chen ZJ. Stochastic and epigenetic changes of gene expression in *Arabidopsis* polyploids. Genetics. 2004.167(4):1961-73.

This dissertation was typed by Jinsuk Lee.